Metabolic diversity in commensal protists regulates intestinal immunity and trans-kingdom competition

Elias R. Gerrick[1], Soumaya Zlitni[2,3,9], Patrick T. West[2,3,9], Matthew M. Carter[4,9], Claire M. Mechler[1], Matthew R. Olm[4], Elisa B. Caffrey[4], Jessica A. Li[1], Steven K. Higginbottom[4], Christopher J. Severyn[2,5], Frauke Kracke[6], Alfred M. Spormann[6,7], Justin L. Sonnenburg[4,8], Ami S. Bhatt[2,3], Michael R. Howitt[1,4,10,*]

[1]Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA

[2]Department of Genetics, Stanford University, Stanford, CA 94305, USA

[3]Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

[4]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA

[5]Department of Pediatrics, (Division of Hematology/Oncology/Stem Cell Transplant and Regenerative Medicine), Stanford University, Palo Alto, CA , 94305, USA

[6]Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA

[7]Department of Chemical Engineering, Stanford University, Stanford, CA 94305, USA

[8]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

[9]These authors contributed equally

[10]Lead contact

*Correspondence: mhowitt@stanford.edu

Keywords

Commensal protist; Tritrichomonas; parabasalid; microbiome; trans-kingdom; mucus; fiber; microbiota-accessible carbohydrate; metabolites; tuft cell

Summary

The microbiota influences intestinal health and physiology, yet the contributions of commensal protists to the gut environment have been largely overlooked. Here, we identified several new rodent- and human-associated parabasalid protists. Genomic and metabolomic analyses of murine parabasalids from the genus *Tritrichomonas* revealed species-level differences in the excretion of the metabolite succinate. This metabolic dissimilarity results in distinct small intestinal immune responses during protist colonization. Metabolic differences between *Tritrichomonas* species also determine their ecological niche within the microbiota. By manipulating dietary fibers and developing *in vitro* protist culture, we show that different parabasalid species preferentially rely on dietary polysaccharides or mucus glycans. These polysaccharide preferences create trans-kingdom competition with specific commensal bacteria, which affects intestinal immunity in a diet-dependent manner. Our findings reveal unappreciated diversity in commensal parabasalids, elucidate differences in commensal protist metabolism, and suggest how dietary interventions could regulate their impact on gut health.

Introduction

The gut microbiota is a complex and diverse microbial ecosystem that broadly impacts the development and function of the intestinal immune system[1–5]. This relationship has been studied predominantly with microbial communities, where conserved microbial ligands and metabolites have pleiotropic effects on the host[6–10]. However, some microbial species dominantly influence the immune system independent of other microbiota members[11–13]. For example, segmented filamentous bacteria induce T helper 17 (Th17) cells in the murine small intestine, and *Bacteroides fragilis* stimulates regulatory T cells in the colon[12,13]. While bacteria are the most well-characterized immunologically dominant symbionts, the gut microbiota also contains viruses, fungi, archaea, and protists that interact with the host and alter immune function.

Protists are often overlooked as members of the microbiota, in part due to the historical view of these microbes as universally pathogenic. However, commensal protists, including those in the *Parabasalia* phylum, are present in healthy mammalian guts and can dominantly shape intestinal immunity in mice[14–18]. Parabasalid protists from the genus *Tritrichomonas* are widespread in laboratory mice and dramatically alter the gut mucosa[15,19–21]. In the small intestine, *Tritrichomonas* species drive type 2 immunity through production of the metabolite succinate, which lengthens the gut and promotes tuft and goblet cell hyperplasia in the epithelium[20,21]. In the colon, *Tritrichomonas musculis* (*Tmu*) activates the inflammasome to promote Th1 and Th17 immunity[15,22]. The effects of these tritrichomonads on epithelial and immune functions significantly alter the severity of some enteric infections and several inflammatory diseases in mouse models, and do not induce apparent pathology[15,20–22]. In humans, only two species of commensal parabasalids have been identified, although their effects on intestinal immunity are still unclear. Thus, despite the growing appreciation of how commensal parabasalids impact mucosal tissues, very little is understood about their diversity, metabolism, or microbial ecology.

77 Fundamental aspects of commensal parabasalid protist biology have been obscured by
78 a lack of genomic information and their resistance to *in vitro* culture. Consequently, the biology
79 of commensal *Tritrichomonas* species is largely inferred from studies on the urogenital
80 pathogenic parabasalids *Trichomonas vaginalis* and *Tritrichomonas foetus.* However, these
81 pathogens infect different hosts and anatomical niches that may not reflect the specific
82 adaptations of commensal parabasalids in the gut. Additionally, the ecological role of
83 commensal parabasalids within the mammalian gut microbiota is unexplored. Parabasalid
84 protists are generally fermentative[21,23–25], but the specific carbohydrate sources used by
85 commensal species and how their harvesting of these carbohydrates affects the intestinal
86 environment is unknown. Thus, parabasalid metabolism may contain unidentified diversity that
87 shapes commensal protists' interactions with both the host and microbiota.
88 In this work, we identified parabasalids in humans from both industrialized and non-
89 industrialized populations, revealing large amounts of unappreciated diversity in human-
90 associated parabasalids. In addition, we identified a novel murine *Tritrichomonas* species called
91 *Tritrichomonas casperi.* Using genomic, transcriptomic, and metabolomic approaches, we
92 determined that fermentative output from *T. casperi* differs from previously characterized murine
93 tritrichomonads like *Tmu,* which results in divergent immune responses in the small intestine.
94 Further, we used this metabolic information to create a framework for predicting the production
95 of immunomodulatory metabolites by parabasalids, including human-associated species, based
96 only on phylogeny. Next, we demonstrated that these different species' preferences for
97 fermenting plant polysaccharides or mucus glycans affect trans-kingdom interactions with
98 bacteria. Finally, through highly controlled studies on carbohydrate utilization, we revealed that
99 the composition of dietary fiber acts as a lever to control protist-induced type 2 immunity. Our
100 work provides mechanistic insights into how commensal parabasalids, through their metabolic
101 differences, act as architects of the intestinal ecosystem and establishes a foundation for further
102 mechanistic studies of how these protists influence host immunity and microbiome ecology.

103

104 **RESULTS**

105 **Discovery of a new commensal *Tritrichomonas* species in mice**

106 To discover uncharacterized parabasalid protists within laboratory mice, we screened
107 DNA extracted from stool using pan-parabasalid PCR primers. These primers bind to highly
108 conserved sequences of parabasalid rRNA loci and generate internally transcribed spacer (ITS)
109 amplicons sufficient to distinguish parabasalid species using Sanger Sequencing. As expected,
110 we identified *Tmu* (previously referred to as *T. muris*[19]) in a subset of our mice, but stool from
111 another group of mice produced mixed signals that indicated the presence of multiple closely
112 related parabasalid species. Examination of the cecal contents of these mice by light
113 microscopy revealed protists with two distinct sizes: a large variant consistent in size with *Tmu*
114 and another variant approximately half the length.

115 To determine if these two sizes constitute different morphologies of *Tmu* or two separate
116 species, we purified both sizes of protist by fluorescence-activated cell sorting (FACS) and
117 colonized protist-negative specific pathogen free (SPF) mice with either the large or small
118 variant (Figure 1A). Mice maintained colonization with homogenously small or large protists, as
119 verified by scanning electron microscopy (SEM), suggesting these were separate species
120 (Figure 1B and S1A). Ultrastructural examination of both protists revealed an undulating
121 membrane and recurrent flagellum typical of the *Parabasalia* phylum (Figure 1B)[26]. Additionally,
122 the new protist, like *Tmu,* had three anterior flagella, suggesting it belongs to the genus
123 *Tritrichomonas*. To confirm this phylogenetic relationship, we compared ITS sequences from
124 both protists, which revealed the smaller species is an uncharacterized murine isolate with
125 accession number MF375342 that shares 83.4% identity with *Tmu* (Figure S1B). Thus, the
126 smaller protist is a new species, hereafter referred to as *Tritrichomonas casperi* (*Tc*). Healthy
127 mice maintained high levels of *Tc* colonization for at least 3 months, indicating that *Tc* is a
128 stable member of the murine microbiota (Figure S1C).

129

130 **Characterization of parabasalid diversity in humans**

131 Rodent *Tritrichomonas* species are not known to colonize humans, but two commensal
132 parabasalid species, *Pentatrichomonas hominis* and *Dientamoeba fragilis*, can be found in the
133 human gut. However, most studies have used targeted approaches that are unlikely to detect
134 diverse species and primarily investigate samples from industrialized populations which harbor
135 less diverse microbiomes[18,27,28]. Therefore, to identify human-associated commensal
136 parabasalids we created a reference set of 29 parabasalid ITS sequences (Table S1) to query
137 previously published metagenomes from 1,800 human fecal samples representing 29 different
138 populations[28]. These metagenomes encompass a wide array of human lifestyles, including
139 industrial and hunter-gatherer populations, as well as "transitional" populations which have
140 neither hunter-gatherer nor industrialized lifestyles. Using our parabasalid ITS database as
141 reference sequences, we assembled ITS sequences from five parabasalids in these human
142 samples, as well as a novel *Simplicimonas* species from cow feces which we refer to as
143 *Tritrichomonadida*_sp_1012. While the human-associated sequences included *D. fragilis* and *P.*
144 *hominis*, they also contained *Tetratrichomonas buttreyi*, which has only previously been
145 identified in cattle. We also identified two previously undescribed candidate parabasalid species
146 in the microbiomes of non-industrialized populations, which we refer to as
147 *Trichomonadida*_sp_97 and *Parabasalia*_sp_1653 (Figure 1C and Table S1). Phylogenetic
148 analysis suggested that *Trichomonadida*_sp_97 belongs to the *Trichomonadida* order and
149 *Parabasalia*_sp_1653 potentially represents a novel *Parabasalia* order (Figure 1C).

150 The method used to identify these six parabasalids in the metagenomic assemblies
151 suffers from low sensitivity inherent to assembly-based approaches. This assembly-based
152 sensitivity issue is further compounded by low sequencing depth in a subset of the
153 metagenomic datasets which may have limited the number of identified protists. We therefore
154 analyzed these datasets using a mapping-based approach, which improves sensitivity of

155 detection but does not enable accurate classification beyond identifying the closest match in our
156 reference database (Figure S1D and 1E). This mapping-based methodology revealed a strong
157 effect of sequencing depth on parabasalid identification; we therefore implemented a 5Gbp
158 minimum sequencing depth threshold for further analysis (Figure S1F). In this curated database
159 of 557 metagenomes, there was an inverse relationship between industrialization and
160 parabasalid colonization frequency, with parabasalids identified in more than 22% of samples
161 from non-industrialized populations but only 3.8% of industrialized samples (Figure 1D, 1E, 1F).
162 *D. fragilis* was the most prevalent parabasalid in industrialized and transitional populations, but
163 not in the hunter-gatherer populations. The microbiomes of non-industrialized individuals
164 contained many other parabasalid species including *T. buttreyi*, *Trichomonadida*_sp_97, and
165 *Parabasalia*_sp_1653. Similarly, parabasalids closely related to *T. vaginalis*,
166 *Tritrichomonadida*_sp_1012, and the murine tritrichomonads *Tmu* and *Tc,* were also present in
167 some non-industrialized populations (Figure 1D, 1E, and 1F). Together, these results
168 demonstrate that a large amount of unappreciated diversity exists among human-associated
169 commensal parabasalids and are consistent with previous studies suggesting that
170 industrialization is associated with decreased diversity of microeukaryotes in the
171 microbiome[17,18].
172

173 **_Tmu_ and _Tc_ stimulate divergent immune responses in the small intestine**

174 *D. fragilis* belongs to the same phylogenetic order as the murine tritrichomonads *Tmu*
175 and *Tc*, and we identified even closer relatives of the murine protists in non-industrialized
176 human populations. Thus, our results reveal new parabasalid diversity within human
177 microbiomes and reinforce the utility of murine tritrichomonads for understanding the functional
178 impact of commensal parabasalids on the intestinal environment. All previously characterized
179 murine commensal tritrichomonads, including *Tmu*, induce tuft cell hyperplasia and type 2
180 immunity in the distal small intestine (SI) through excretion of the metabolite succinate[19–22].

181 However, tuft cell hyperplasia, which is used as a readout for SI type 2 immunity[19–21], was
182 absent in mice colonized with *Tc* (Figure 2A, 2B, 2C, and S2A). Although *Tc* stimulated a subtle
183 increase in group 2 innate lymphoid cells (ILC2s), which are more sensitive to weak type 2
184 stimuli, this increase was markedly lower than the ILC2 expansion associated with *Tmu* (Figure
185 2D, 2E, and S2B). Moreover, *Tc* colonization reduced the proportion of small intestinal GATA3^{+}
186 CD4^{+} T helper 2 (Th2) cells associated with adaptive type 2 immunity, which contrasted with the
187 expansion of Th2 cells in *Tmu* colonized mice (Figure 2F and S2C). To determine whether *Tc*
188 failed to drive type 2 immunity because of colonization deficiencies, we compared *Tc* and *Tmu*
189 abundance by quantitative PCR. *Tc* was at least as abundant as *Tmu* along the length of the
190 intestine, including the distal SI (Figure S2D). These results suggest that *Tc* is unique among
191 known murine commensal tritrichomonads because it does not induce type 2 immunity in the
192 distal SI.

193 Previous work demonstrated that *Tmu* increases CD4^{+} T helper 1 (Th1) and T helper 17
194 (Th17) cells in the cecum and colon[15,22]. Therefore, we examined whether *Tc* stimulates these
195 same immune responses in the distal SI. IFNγ producing Th1 cells and IL-17 producing Th17
196 cells, as well as IFNγ and IL17 double positive CD4^{+} T cells, significantly increased in the distal
197 SI of *Tc* colonized mice compared to protist-free controls (Figure 2G, 2H, and S2C). Because
198 *Tmu* stimulates Th2 cells in the distal SI, we wondered whether co-colonization would alter the
199 mucosal response to these protists. To address this question, we co-colonized mice with *Tmu*
200 and *Tc* or singly with each protist. Tuft cell hyperplasia was identical in co-colonized and *Tmu*
201 singly colonized mice, and *Tc* and *Tmu* abundance was unaffected by co-colonization (Figure
202 S2E and S2F). This finding suggests that *Tmu* dominates the mucosal immune response when
203 both species are present simultaneously in the distal SI.
204

205 **Th1 and Th17 induction is a common tritrichomonad immune response**

206 Although *Tmu* induces Th1 and Th17 cells in the colon and type 2 immunity in the distal
207 SI, no study has examined the immune response in both intestinal tissues using the same
208 protist isolate. This leaves some ambiguity regarding whether different strains or species
209 contribute to these heterogeneous *Tmu*-induced immune responses. We therefore
210 characterized $CD4^+$ T cell populations induced by *Tmu* and *Tc* in the colon. Both *Tmu* and *Tc*
211 significantly increased Th1, Th17, and IFNγ-IL17 double positive T cells in SPF and gnotobiotic
212 mice (Figure 3A, 3B, and S3A). Other T cell responses appeared unaffected by protist
213 colonization; for example, regulatory T cells did not change with colonization by either species
214 (Figure S3B).
215 Microbial adhesion has been proposed as a nonspecific Th17-inducing signal[29], and
216 previous work identified examples of *Tmu* adhering to colonic epithelial cells[15]. To address the
217 prevalence of epithelial adhesion by *Tmu* and *Tc,* we developed a polyclonal antibody against
218 *Tmu*. This antibody recognizes *Tmu,* and to a lesser extent *Tc,* and stains both trophozoites and
219 pseudocysts of *Tmu* (Figure S3C and S3D). We gently flushed the colons of *Tmu* colonized
220 mice with PBS to remove unadhered protists and then searched for adherent protists by
221 fluorescence microscopy. Adherent protists were extremely rare (0-3 per colonic tissue section)
222 and thus are unlikely to drive immune responses across the entire tissue (Figure S3E, S3F, and
223 S3G).
224 *Tmu* requires the inflammasome to fully induce colonic Th1 and Th17 cells[15], so we
225 hypothesized that *Tc* also stimulates Th1 and Th17 immunity by activating the inflammasome.
226 Mice deficient in the inflammasome component Caspase 1 had partially suppressed colonic
227 immune induction by *Tmu* (Figure 3A and 3B)[15]. Caspase 1 was similarly important for the
228 immune response in *Tc* colonized mice, suggesting that both protists stimulate Th1 and Th17
229 immunity in the colon through a shared effector.
230 Although *Tmu* and *Tc* induce the same Th1 and Th17 immune responses in the colon,
231 only *Tc* appears to drive this response in the distal SI. Induction of Th1 and Th17 immunity in

232 the distal SI by *Tc* demonstrates that the host pathway is present. Therefore, we reasoned that
233 the lack of Th1 and Th17 induction in the distal SI by *Tmu* may be due to succinate driven type
234 2 immunity masking this response. The poor cellular viability caused by processing tissue with
235 active type 2 immunity prevented direct measurement of IFNγ^{+} and IL17^{+} T cells in the distal SI
236 of wild type *Tmu* colonized mice. Instead, we measured Th1 and Th17 induction in *Trpm5*$^{-/-}$
237 mice, which lack the tuft cell taste chemosensory component TRPM5 and are insensitive to
238 succinate-induced type 2 immunity[20,21]. Like *Tc* in wild type mice, *Tmu* stimulated expansion of
239 IFNγ^{+}, IL17^{+}, and IFNγ-IL17 double positive CD4^{+} T cells in the distal SI of *Trpm5*$^{-/-}$ mice (Figure
240 3C and 3D). Altogether, these data suggest tritrichomonads can induce Th1 and Th17 immunity
241 across the intestinal tract, but metabolic production of succinate by *Tmu* masks this response in
242 the distal SI by driving type 2 immunity.
243
244 **Metabolic differences in tritrichomonads underly the divergent immune responses**
245 Succinate excreted by *Tmu* stimulates the succinate receptor (GPR91) on tuft cells and
246 drives type 2 immunity in the distal SI (Figure S4A)[20,21]. Because *Tc* does not stimulate type 2
247 immunity, we hypothesized that *Tc* does not produce succinate. We directly measured *in vivo*
248 succinate concentrations in the ceca of *Tmu* and *Tc* colonized mice using liquid
249 chromatography-mass spectrometry (LC-MS). *Tmu* colonization significantly increased the
250 extracellular concentrations of cecal succinate compared to uncolonized controls, while *Tc*
251 colonized mice showed no change in succinate levels (Figure 4A).
252 Two fermentative pathways are currently established in parabasalid protists based on
253 studies of *T. vaginalis* and *T. foetus* metabolism[24,30]. *T. vaginalis* redox homeostasis converts
254 pyruvate to lactate through the action of lactate dehydrogenase. In *T. foetus,* and presumably
255 *Tmu*, redox balance occurs via conversion of malate to succinate by the sequential activity of
256 fumarase and fumarate reductase. Lactate or succinate are then excreted as waste products

257 from the protists. However, as with succinate, *Tc* colonization did not result in an increase in
258 extracellular lactate in mouse cecal contents (Figure S4B).

259 To identify the metabolic pathways present in each protist, we sequenced the genomes
260 of both *Tmu* and *Tc*. Massive numbers of repeat regions in parabasalid genomes make short-
261 read based genome assemblies challenging[31–33]. Therefore, we used Nanopore long-read
262 sequencing and short-read error correction to assemble these parabasalid genomes (Figure 4B,
263 4C, and S4C). As predicted, *Tmu* encodes the succinate production pathway, but not the lactate
264 pathway (Figure 4D, cyan box). Despite the lack of succinate or lactate excretion by *Tc*, it
265 encodes both the succinate and lactate production genes (Figure 4D, magenta boxes). To
266 assess the relative contribution of each pathway to *Tc* metabolism, we performed RNA
267 sequencing on *Tmu* and *Tc* isolated from the distal SI and ceca of mice. *Tmu* and *Tc* both
268 expressed genes in the succinate pathway at high levels, whereas lactate dehydrogenase was
269 less highly expressed in *Tc* (Figure 4E and 4F). Transcriptomics also confirmed that expression
270 of these fermentative pathways does not vary drastically between the distal SI and cecum.
271 Together, these data suggest that *Tmu* and *Tc* both use the succinate fermentation pathway,
272 contrary to our hypothesis based on LC-MS measurements from cecal supernatants and the
273 lack of type 2 immune induction by *Tc*.

274 To directly determine the usage of each fermentative pathway, we measured the
275 intracellular concentrations of both metabolites in lysates from purified protists. Consistent with
276 the genomic and transcriptomic analyses, cell extracts from purified *Tmu* and *Tc* contained high
277 levels of succinate and low levels of lactate (Figure 4G and 4H). These intracellular succinate
278 concentrations suggest that both protists use the succinate pathway to maintain redox
279 homeostasis, yet cecal LC-MS measurements and the absence of tuft cell hyperplasia suggest
280 *Tc* does not excrete this metabolite. Some bacteria can metabolize succinate into short chain
281 fatty acids such as propionate and butyrate[34,35]. Therefore, we reasoned that *Tc* might also
282 metabolize succinate in a similar manner. However, short chain fatty acids, including butyrate,

283 propionate, and acetate, were not more abundant in lysates from *Tc* than *Tmu* (Fig. S4D). We
284 next attempted to identify other potential succinate derivatives in *Tc* using additional mass
285 spectrometry methods capable of identifying over 400 microbiota-produced metabolites[36], but of
286 the 48 metabolites we identified none had significantly higher abundance in *Tc* (Figure S4E).
287 Altogether, these data indicate that although both protists use succinate production for redox
288 homeostasis *Tc* does not excrete sufficient succinate into the intestinal lumen to stimulate type
289 2 immunity. However, the precise metabolic mechanism remains undefined.
290 We observed that *T. foetus*, *Tmu*, and *Tc* all use the succinate fermentative pathway and
291 belong to the order *Tritrichomonadida*, whereas *T. vaginalis* uses the lactate pathway and
292 belongs to the order *Trichomonadida*. This led us to hypothesize that an evolutionary split had
293 occurred within the *Parabasalia* phylum, with lactate versus succinate production representing a
294 branch point. To investigate this hypothesis, we analyzed the fermentative metabolism of two
295 additional intestinal parabasalids: *P. hominis* (a human-associated *Trichomonadida*) and
296 *Monocercomonas colubrorum* (a reptile-associated *Tritrichomonadida*). As predicted, *P. hominis*
297 cultures contained high concentrations of lactate whereas *M. colubrorum* cultures contained
298 high concentrations of succinate (Figure S4F). These data suggest that the phylogenetic
299 classification of new and uncharacterized parabasalid species, such as those identified in this
300 study, can help predict the production of the immunomodulatory metabolites lactate and
301 succinate (Figure 4I). For example, this predictive framework suggests that the human-
302 associated parabasalid *Trichomonadida*_sp_97 produces lactate, whereas the novel human-
303 associated *Tritrichomonas* species likely use the succinate fermentative pathway (Figure 4I).
304

305 **_Tmu_ and _Tc_ occupy different nutritional niches within the microbiota**

306 In addition to the microbiota's metabolic output, the metabolic input of each species also
307 shapes its interactions with other microbes and the host[37–39]. Plant polysaccharides such as
308 dietary fibers are the most commonly utilized carbon source by the microbiota, and collectively

309 are referred to as dietary Microbiota Accessible Carbohydrates (MACs)[40–42]. However, non-
310 dietary MACs are also prevalent in the intestine. For example, mucus secreted by intestinal
311 epithelial goblet cells provides a complex and glycan-rich nutrient source for microbes capable
312 of breaking down its diverse glycosidic linkages[42–44]. To determine whether *Tmu* and *Tc* ferment
313 dietary MACs, we altered the fiber content of mouse diets using defined chow similar to a
314 previous study[21]. Fibers in these diets were restricted to inulin, a fermentable fiber that acts as a
315 dietary MAC, and cellulose, a non-fermentable fiber that acts as a bulking agent and passes
316 largely undigested through the intestine. Standard mouse chow, which contains a complex
317 mixture of plant material, was used as a control (complex). *Tmu* abundance was high in mice
318 fed complex chow and defined chow with inulin and cellulose (def-CI) or inulin alone (def-I)
319 (Figures 5A and S5A). However, *Tmu* colonization was dramatically reduced when fiber was
320 removed entirely (def-None) or cellulose was the sole fiber (def-C) (Figures 5A and S5A). In
321 contrast, *Tc* colonization was completely independent of dietary MACs, as this protist's
322 abundance did not decrease on any of the defined diets (Figure 5B, S5B). These results
323 suggest that *Tc* may rely on host-derived carbohydrates as a nutritional source, while *Tmu*
324 colonization depends on dietary MACs.

325 The gut microbial ecosystem is a complex and species rich community, with most
326 nutritional niches occupied by one or more community members competing for limited
327 resources[37,45,46]. This scarcity of open niches can prevent new microbes from colonizing the gut
328 and establishing themselves in the community, a phenomenon termed "colonization resistance".
329 However, commensal tritrichomonads such as *Tmu* and *Tc* are unusual in that they appear
330 unaffected by colonization resistance. As few as one thousand *Tmu* or *Tc* protists can stably
331 colonize SPF mice without niche-clearing manipulation. We reasoned that *Tmu* and *Tc*
332 overcome colonization resistance due to one of two possibilities, (1) *Tmu* and *Tc* occupy unique
333 and uninhabited niches within the gut ecosystem, or (2) *Tmu* and *Tc* outcompete other microbes
334 that occupy their preferred niche. Due to the dearth of unoccupied niches, we favored the

335 second hypothesis. Therefore, identifying microbes eliminated by tritrichomonad colonization
336 could offer insights into the preferred nutritional niches of these protists.
337 To investigate trans-kingdom competition between tritrichomonads and bacteria, we
338 performed 16S rRNA sequencing on the distal SI, cecum, and colon contents from mice with
339 and without *Tmu* or *Tc* (Figures 5C, S5C, and S5D). In all three regions of the intestine, *Tmu*
340 colonization dramatically reduced *Bifidobacterium pseudolongum* and *Turicibacter sanguinis*
341 abundance (Figure 5C, 5D, S5C, and S5D). Alternatively, *Tc* colonization decreased
342 *Akkermansia muciniphila* to undetectable levels (Figures 5C, 5D, and S5C-E). *Bifidobacteria* are
343 prolific fiber degraders and *Turicibacter* abundance is positively correlated with fiber
344 consumption[47,48], consistent with *Tmu* outcompeting these bacteria for dietary MACs. *A.*
345 *muciniphila* is a mucin specialist, suggesting that *Tc* may preferentially consume host mucus
346 glycans.
347 In the colon, mucus forms a highly interlinked network that is tightly associated with the
348 intestinal epithelium and acts as a barrier against microbial encroachment[2]. Mucolytic bacteria,
349 such as *A. muciniphila*, preferentially colonize the mucus layer. To determine whether *Tc* also
350 localizes to the colonic mucus layer, we used quantitative imaging to map the location of each
351 protist along the mucosal-luminal axis of the intestine. *Tmu* was present at both the mucosal
352 interface and in the lumen of the colon, consistent with a preference for dietary MACs (Figures
353 5E and 5F). *Tc* instead tightly co-localized with the colonic mucus layer, corroborating its
354 preference for mucus glycans.
355 Mucus glycans and fiber are composed of diverse sugar structures that require a large
356 repertoire of carbohydrate active enzymes (CAZymes) to catabolize. To investigate the glycan
357 utilization potential of *Tmu* and *Tc*, we used dbCAN2 to identify CAZymes in the protist
358 genomes[49]. Additionally, we annotated CAZymes in the genomes of the potential bacterial
359 competitors *B. pseudolongum*, *T. sanguinis*, and *A. muciniphila.* Consistent with our hypothesis
360 of nutritional trans-kingdom competition, *B. pseudolongum* and *T. sanguinis* were enriched for

361 CAZymes active on plant polysaccharides, whereas *A. muciniphila* was enriched for host-active
362 CAZymes (Figures 5G and S5F). However, *Tmu* and *Tc* each encoded a large number of both
363 plant and host-active CAZymes, suggesting that these protists are capable of breaking down a
364 wide variety of carbohydrate substrates. We therefore used the protist transcriptomes to
365 determine the expression levels of mucus active CAZymes. *Tc* expresses CAZymes with
366 predicted activity against all of the sugar subunits found on mucus glycans, and these CAZymes
367 are members of Glycoside Hydrolase (GH) families associated with mucus degradation (Figure
368 5H)[27,43]. In particular, *Tc* expresses high levels of galactosidases, which can cleave the
369 abundant galactose residues from mucus glycans. Galactose cannot be directly catabolized and
370 must be converted to glucose to be oxidized for energy production. Consistent with a galactose-
371 dependent metabolism, *Tc* also expresses high levels of UDP-galactose 4-epimerase used to
372 convert galactose to glucose[50]. However, these putatively mucolytic CAZymes identified in *Tc*
373 are also expressed by *Tmu,* albeit at lower levels of expression, suggesting that *Tmu* may also
374 be capable of fermenting mucus glycans (Figure 5H).
375
376 **Dietary MAC deprivation causes *Tmu* to switch to mucolytic metabolism**
377 Although *Tmu* colonization was markedly reduced when mice were fed the def-C diet, a
378 small number of protists survived in a subset of these mice (Figure 5A and S5A). This
379 observation, combined with the expression of mucolytic CAZymes, suggested that *Tmu* can
380 utilize mucus glycans to survive dietary MAC deprivation. To investigate this possibility, we
381 examined the localization of *Tmu* in the colons of mice fed cellulose as the sole fiber. The few
382 remaining *Tmu* cells in def-C fed mice strongly co-localized with the mucus layer (Figure 6A and
383 6B).
384 We next sought to study this metabolic shift in a reductionist model to understand the
385 effects of specific substrates on protist growth and survival without the inherent complexity of
386 the intestinal environment. Although some parabasalid species can be propagated *in vitro*,

387 murine gut commensals such as *Tmu* and *Tc* do not grow with existing culture conditions.
388 Although previous studies by our group and others have cultured these protists short-term to
389 eliminate bacteria and fungi with antibiotics[15,19–21], growth does not occur; instead, the protists
390 slowly die over the course of several days. This method is sufficient to colonize mice with pure
391 protists, but it does not allow for mechanistic investigation of protist biology. We created a new
392 culture medium, which we named PBF, that supports the long-term culture of *Tmu* (Figure S6A).
393 Despite the success of PBF for *Tmu* culture, it did not support the growth of *Tc*. Nevertheless,
394 adding mucus to PBF medium improved *Tc* survival 5-fold, supporting the conclusion that this
395 species ferments mucus glycans (Figure S6B).

396 Next, we cultured *Tmu* axenically with either the simple sugar maltose or with mucus as
397 the primary carbon source and found that both glycans supported robust protist growth (Figure
398 6C). To confirm whether the CAZymes expressed by protists *in vivo* (Figure 5H) may support
399 mucus glycan harvesting, we performed RNA sequencing on *Tmu* grown with either maltose or
400 mucus (Figure 6D). *Tmu* grown with mucus significantly increased the expression of 16
401 CAZymes in GH families associated with mucolytic activity relative to protists grown with
402 maltose. These upregulated CAZymes included the putative mucolytic genes identified above,
403 supporting the hypothesis that both protists use these CAZymes to metabolize mucus glycans.
404 In addition, the gene encoding the galactose-utilization enzyme UDP-galactose-4-epimerase
405 was upregulated in mucus-grown *Tmu,* suggesting a metabolic shift towards fermenting
406 mucosal sugars.

407 The discrepancy between *Tmu* survival on mucus *in vivo* (in which the protists barely
408 survive) and *in vitro* (where the protists thrive) led us to consider what factors might account for
409 this difference. One major distinction is the absence of bacteria in the *in vitro* cultures. To
410 determine if mucolytic bacteria outcompete *Tmu* during dietary MAC deprivation *in vivo*, we
411 treated *Tmu*-colonized mice with antibiotics before switching to the def-C diet. Strikingly,
412 antibiotic treatment with a cocktail of vancomycin, neomycin, and ampicillin restored *Tmu*

413 colonization in the absence of dietary MACs (Figure 6E). To narrow down the potential bacterial
414 competitors, we tested the ability of each antibiotic to restore colonization. Neomycin and
415 vancomycin were dispensable for *Tmu* survival, whereas ampicillin treatment completely
416 restored colonization (Figure 6F and S6C). Consistent with ampicillin providing *Tmu*
417 uncontested access to the mucus niche during fiber deprivation, the protists localized tightly to
418 the mucus layer (Figure S6D and S6E).
419 To identify specific bacteria that outcompete *Tmu* during dietary MAC starvation, we
420 profiled the changes to the bacterial microbiome during treatment with each of the individual
421 antibiotics. Mice treated with ampicillin had undetectable levels of *Bacteroidetes* and *A.*
422 *muciniphila*, known mucolytic members of the microbiota (Figure 6G, 6H, and S6F)[43]. In
423 contrast, the mice treated with no antibiotics, vancomycin, or neomycin all had high levels of
424 *Bacteroidetes*, *A. muciniphila*, or both. These results suggest that *Tmu* switches to a mucolytic
425 metabolism during dietary MAC starvation but is largely outcompeted by the mucolytic bacteria
426 *Bacteroidetes spp*. and *A. muciniphila*.
427
428 **Diet-protist interactions shape the host immune response**
429 Next, we characterized dietary MAC fermentation by *Tmu* using axenic *in vitro* culture.
430 Although inulin was sufficient for *Tmu* to colonize mice (Figure 5A), *Tmu* failed to grow in PBF
431 medium with inulin as the carbon source (Figure 7A). To understand this apparent discrepancy,
432 we investigated the localization of *Tmu* in mice fed def-CI chow, in which inulin is the sole
433 dietary MAC. Similar to ampicillin treated mice starved of dietary MACs, protists were tightly
434 associated with the mucus layer in mice fed def-CI chow (Figure 7B and 7C). This suggests that
435 although inulin is sufficient for *Tmu* to colonize mice at high abundance, *Tmu* does not directly
436 utilize inulin and instead ferments mucus glycans when inulin is the sole dietary MAC. Notably,
437 many *Bacteroidetes* species can ferment both inulin and mucus, suggesting that these bacteria
438 may utilize inulin when the fiber is present, thereby reducing competition with *Tmu* for mucus

439 glycans[43,51]. *Bacteroidetes* abundance did not change in *Tmu*-colonized mice fed def-C chow
440 compared to def-CI chow, whereas *Tmu* levels decreased in def-C chow (Figure 7D and 5A).
441 This observation is consistent with *Bacteroidetes* outcompeting *Tmu* for mucus glycans when all
442 dietary MACs, including inulin, are removed. However, this same dietary switch in *Tc*-colonized
443 mice results in a significant decrease in *Bacteroidetes* abundance, consistent with the ability of
444 *Tc* to outcompete mucolytic bacteria for the mucus niche (Figure 7D).

445 To determine whether *Tmu* can ferment other dietary MACs, we tested a wider array of
446 fermentable fibers as the carbon source in *Tmu* cultures. While *Tmu* failed to grow in maltose-
447 free PBF medium supplemented with the fibers arabinan, arabinogalactan, and
448 fructooligosacharides (FOS), the protist readily grew on arabinoxylan, galactomannan, and
449 pectin (Figure 7E and S7A). To identify CAZymes responsible for fiber utilization by *Tmu*, we
450 performed transcriptomics on *Tmu* cultured with either the simple sugar maltose or the
451 fermentable fiber arabinoxylan (Figure 7F). We identified 3 CAZymes in GH families associated
452 with fiber utilization that increased expression during growth on arabinoxylan, suggesting these
453 CAZymes break down fiber polymers[27,43].

454 Finally, we wondered whether diet and *Tmu* metabolic input would affect immune
455 induction by the protist. When *Tmu*-colonized mice were fed def-C chow, and thus protists were
456 largely depleted from the intestine, type 2 immune induction was completely ablated (Figure
457 S7B and S7C). However, when *Tmu*-colonized mice were fed def-CI chow, and thus *Tmu*
458 fermented mucus glycans, type 2 immune induction in the distal SI still occurred but was largely
459 suppressed (Figure 7G and 7H). This suppression did not appear to be due to decreased
460 succinate output by *Tmu* on a per-cell basis during mucus feeding, as *Tmu* grown on mucus *in*
461 *vitro* showed no decrease in succinate excretion (Figure S7D). However, *Tmu* abundance in the
462 distal SI of mice fed def-CI chow was substantially lower than in mice fed complex chow,
463 suggesting that when fermenting mucus, this protist cannot colonize the SI at sufficiently high

464 levels to induce type 2 immunity (Figure 7I). Altogether, these results demonstrate that
465 alteration of dietary fiber intake can induce graded type 2 immune responses by *Tmu*.
466
467 **Discussion**
468 The role of the eukaryotic microbiome, and commensal protists in particular, in shaping
469 community structure and influencing host physiology is far less understood than their bacterial
470 counterparts. This study reveals previously unrecognized phylogenetic diversity of commensal
471 protists in the *Parabasalia* phylum in both mice and humans. Amongst the newly identified
472 human-associated parabasalids are close relatives of the murine species, *Tmu* and *Tc*.
473 Although the species found in humans are distinct from the murine isolates, they reinforce the
474 potential translational impact of functional studies using *Tmu* and *Tc*. We demonstrate
475 numerous shared features between parabasalids, but importantly, our results also revealed that
476 even minor phylogenetic differences can have meaningful biological consequences.
477 Although *Tmu* and *Tc* are closely related species, they differ in the excretion of the
478 fermentative waste product succinate, and they prefer distinct nutritional niches within the gut
479 ecosystem. These metabolic differences lead to divergent small intestinal immune responses,
480 distinct trans-kingdom interactions, and differential reliance on host dietary fiber. For example,
481 succinate induced type 2 immunity in the SI has been a ubiquitous feature of previously
482 characterized commensal tritrichomonads[19–21], but *Tc* demonstrates that this is not universal.
483 Additionally, all commensal parabasalids are assumed to ferment dietary fiber, but *Tmu* and *Tc*
484 diverge in dietary fiber usage, leading to trans-kingdom competition with fiber or mucus
485 digesting bacteria, respectively. However, comparing *Tmu* and *Tc* revealed unanticipated
486 similarities as well. *Tc* increases Th1 and Th17 cells in the distal SI, whereas *Tmu* initiates type
487 2 immunity by stimulating tuft cells with succinate. However, *Tmu* colonization of mice lacking
488 the tuft cell response to this metabolite instead exhibited a Th1 and Th17 response in the SI
489 comparable to *Tc*. This result suggests both protists have the capacity to stimulate Th1 and

490 Th17 immunity in the small intestine. Additional studies will be required to identify potential
491 shared factor(s) that promote this immune response and determine the extent of its
492 conservation in the *Parabasalia* phylum.

493 Our studies generated new tools and datasets for commensal tritrichomonads, which will
494 eliminate critical barriers that have stymied mechanistic studies of these protists. The genomes
495 of *Tmu* and *Tc* combined with metabolomic and transcriptomic datasets form a valuable
496 foundation for future studies. Furthermore, the ability to grow these tritrichomonads in culture is
497 an essential step for microbiological investigations. Previous studies, including our own, isolated
498 tritrichomonads from mice, but the protists did not grow robustly in axenic *in vitro* culture[15,19–21].
499 This basic capability is particularly crucial for members of the microbiota because the complexity
500 of their natural ecosystem confounds the ability to link phenotypes to single members of the
501 community. In addition, the high degree of strain-level variation and potential for genetic drift
502 make growth in culture and the ability to freeze/thaw isolates invaluable for increasing
503 experimental rigor. Further, culture represents the first major hurdle for developing any microbe
504 into a genetically tractable model.

505 Only two species of commensal parabasalid, *D. fragilis* and *P. hominis*, have been
506 identified in the human gut. Previous studies also suggested industrialized microbiomes have
507 decreased microeukaryotic prevalence[17,18]. Here, we identified multiple new human-associated
508 commensal parabasalids, which were primarily found in the stool of individuals from non-
509 industrialized populations. These new parabasalids further emphasize the need to consider the
510 microbiome of diverse human populations and highlight the importance of understanding how
511 different parabasalids influence human health. Our discovery that parabasalid production of the
512 immunomodulatory metabolites succinate and lactate can be predicted based solely on
513 phylogeny is a powerful tool. For example, this model predicts that *D. fragilis* and the newly
514 discovered human-associated tritrichomonads produce succinate, and therefore may stimulate
515 type 2 immunity in the small intestine. However, further work is required to refine this predictive

516 model to account for alternative metabolic mechanisms, including those that may limit succinate
517 excretion in protists like *Tc*.
518
519 **Limitations of the study**
520 An important limitation of this study is that our identification of parabasalids in human
521 microbiomes likely includes a high rate of false negativity. Despite the high sequencing depth in
522 some of the metagenomic datasets, read depth was low for parabasalid ITS sequences in most
523 samples. Thus, the prevalence of human-associated parabasalids identified here likely
524 underestimates the true prevalence. Parabasalid-specific approaches will help to circumvent
525 this issue in future studies. A second limitation of this study is that our predictive framework for
526 fermentative pathways in parabasalids only allows for prediction of utilization of a pathway for
527 redox homeostasis, and not excretion of the metabolite. Our discovery that *Tc* produces
528 succinate but does not excrete sufficient quantities to stimulate type 2 immunity highlights the
529 importance of this distinction. Identification of this metabolic mechanism will facilitate the
530 prediction of excretion based solely on genomic data.
531
532 **Acknowledgements**
533 We thank members of the Howitt lab, John Boothroyd, Markus Lakemeyer, Rebecca Gellman,
534 and Fatima Enam for their helpful discussions and expertise. We also thank Dylan Maghini, Eli
535 Moss, Roshni Patel, and Sandra Kong for their NGS expertise. We thank John Perrino and Ruth
536 Yamawaki at the Stanford Cell Sciences Imaging Facility for expertise and help with sample
537 processing and scanning electron microscopy. We thank Dr. R. Margolskee, Dr. J. von Moltke,
538 and Amgen for kindly providing the *Trpm5*$^{-/-}$, SMART13, and *Gpr91*$^{-/-}$ mice, respectively. This
539 work was supported by grants from the NIH (R01DK128292, K01DK113041, R01AI14862302,
540 R01AI14375702, 5 T32 AI07290, and T32HL120824), a Jacob Churg award, and Stanford
541 Pediatric IBD and Celiac Disease Research Program and the Stanford Maternal & Child Health

Research Institute Seed Grant to M.R.H.. E.R.G. is supported by the Stanford Pediatric IBD and Celiac Disease Research Program and the Stanford Maternal & Child Health Research Institute.

Author contributions

E.R.G, S.Z., P.T.W., M.M.C, M.R.O, A.S.B., and M.R.H. designed the experiments. E.R.G., S.Z., P.T.W., M.M.C., M.R.O., E.B.C., C.M.M., J.A.L., S.K.H. performed experiments and contributed intellectual expertise. J.L.S. and A.S.B., provided intellectual expertise. E.R.G. and M.R.H. conceived the study and wrote the manuscript, and all authors read and commented on the manuscript.

Declaration of interests

The authors declare no relevant competing interests.

568 **Figure legends**

569 **Figure 1. Identification of novel parabasalids in mice and humans.**

570 (A) FACS purification of the large *T. musculis* (*Tmu*) and small *T. casperi* (*Tc*) protists isolated
571 from a mouse cecum. FSC, forward scatter. SSC, side scatter. (B) SEM images of *Tmu* (left)
572 and *Tc* (right). UM: Undulating Membrane, RF: Recurrent Flagellum, AF: Anterior Flagella.
573 Scale bars are 2μm. (C) Cladogram of parabasalid protists based on ITS sequences, including
574 novel mouse and human-associated parabasalids. Protists in bold indicate novel parabasalids
575 discovered in this study. (D-F) Identification of human-associated parabasalids using mapping-
576 based analysis of metagenomic data. (D) Percentage average nucleotide identity (ANI) of
577 parabasalid metagenomic reads mapping to the corresponding reference protist sequence.
578 Imperfect matches indicate that identified parabasalids are relatives of reference protists and
579 not the same species. First author of the cohort study, country of origin, and population type
580 ("Industrial", "Transitional", Hunter-Gatherer ("H-G")) is indicated. Columns and rows are
581 hierarchically clustered based on Euclidean distance. Cells corresponding to protists not
582 identified in a population are colored grey. (E) Percentage prevalence of subjects in human
583 cohorts with metagenomic reads mapping to reference parabasalids. (F) Prevalence of
584 parabasalids identified in each population type. Asterisks denote protists with imperfect matches
585 to the reference sequence, which are therefore relatives of the indicated species (E, F).
586 (Abbreviations: TZA=Tanzania, NEP=Nepal, MNG=Mongolia, USA=United States of America,
587 MDG=Madagascar, KAZ=Kazakhstan, SWE=Sweden, HMP=Human Microbiome Project).

588

589 **Figure 2. T Cell responses to *Tmu* and *Tc* differ in the distal SI.**

590 (A) Tuft cell frequency measured by flow cytometry in the distal SI epithelium of uncolonized
591 (*Ctrl*) mice or those colonized with *Tmu* or *Tc* for two weeks. (B) Representative flow cytometry
592 plots of tuft cells. SiglecF marks tuft cells, EpCAM marks epithelial cells. (C) Representative
593 immunofluorescence microscopy images of tuft cell hyperplasia in the distal SI of mice with or

594 without each protist for two weeks. Scale bars are 50μm. (D) Frequency of ILC2s in the lamina
595 propria of the distal SI in *Ctrl* mice or mice colonized with *Tmu* or *Tc* for two weeks. IL13 was
596 measured using the IL13 reporter Smart13 mice (E) Representative flow cytometry plots of
597 ILC2s. (F) Frequency of GATA3^{+} CD4^{+} T cells (Th2 cells) in the lamina propria of the distal SI
598 after three weeks of colonization. (G) Frequency of IFNγ and IL-17 positive CD4^{+} T cells in the
599 distal SI lamina propria after three weeks of *Tc* colonization, or without protist colonization (*Ctrl*).
600 Cytokines were stimulated *ex vivo* and cells were stained for intracellular cytokines. (H)
601 Representative flow cytometry plots of Th1 and Th17 cells in the distal SI with and without *Tc*.
602 Graphs (A, D, F, G) depict mean. Each symbol represents an individual mouse from two to
603 three pooled experiments (A, D, F, and G). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, NS,
604 not significant by Student's t test.
605
606 **Figure 3. Th1 and Th17 induction is the shared tritrichomonad-induced immune**
607 **response.**
608 (A) Frequency (top) and absolute abundance (bottom) of IFNγ and IL-17 positive CD4^{+} T cells in
609 the colonic lamina propria of WT and *Caspase1*$^{-/-}$ mice after three weeks of protist colonization,
610 or without protist colonization (*Ctrl*). (B) Representative flow cytometry plots of the frequency of
611 IFNγ and IL17 positivity in colonic lamina propria CD4^{+} T cells. (C) Frequency and absolute
612 abundance of IFNγ and IL-17 positive CD4^{+} T cells in the distal SI lamina propria of *Trpm5*$^{-/-}$
613 mice after three weeks of *Tmu* colonization, or without protist colonization (*Ctrl*). (D)
614 Representative flow cytometry plots of the frequency of IFNγ and IL17 positivity in distal SI
615 lamina propria CD4^{+} T cells from *Trpm5*$^{-/-}$ mice with or without *Tmu*. Cytokines were stimulated
616 *ex vivo* and cells were stained for intracellular cytokines. Graphs depict mean. Each symbol
617 represents an individual mouse from two (C) or three (A) pooled experiments. *p<0.05,
618 **p<0.01, ***p<0.001, ****p<0.0001, NS, not significant by Student's t test.
619

620 **Figure 4. Metabolic differences in tritrichomonads underly divergent immune responses.**
621 (A) Extracellular succinate concentration in the cecal contents of uncolonized mice (*Ctrl*) or mice
622 colonized with each protist for 3 weeks. (B, C) Genome assemblies of *Tmu* (B) and *Tc* (C).
623 Outermost rings show position (in Mbp), middle ring shows genome assembly contigs, inner ring
624 shows GC content in 10kb sliding windows (blue, above genome average; red, below genome
625 average). Total genome size is at the center of each circle. (D) Schematic of metabolic
626 pathways in parabasalid protists. Magenta boxes surround fermentative pathways found in *Tc*,
627 cyan box surrounds the pathway found in *Tmu*. (E and F) Gene expression levels (in transcripts
628 per million (TPM)) in the cecum vs. distal SI for *Tmu* (E) and *Tc* (F). Transcripts in the lactate
629 and succinate fermentative pathways are labelled if present. (G and H) Intracellular succinate
630 (G) and lactate (H) measurements from purified *Tmu* and *Tc* cell lysates. (I) Cladogram of
631 parabasalids. Green shading shows predicted succinate producers. Purple shading shows
632 predicted lactate producers. Asterisks signify new metabolic confirmations in this study. "Novel
633 human *Tritrichomonadida/Trichomonadida*" represent novel human-associated parabasalids
634 detected in Figure 1, but for which sequence information was not obtainable. Phylogenetic
635 orders are labelled on the right. Graphs (A, G, H) depict mean with SD. *$p<0.05$, **$p<0.01$, NS,
636 not significant by Student's t test.
637
638 **Figure 5. *Tmu* and *Tc* occupy different nutritional niches within the microbiota.**
639 (A) Abundance of *Tmu* and (B) *Tc* in the stool of mice fed diets with defined fiber compositions
640 for two weeks or fed standard chow (complex) as a control. Fibers in the defined diets are
641 limited to inulin (I), cellulose (C), or no fiber (None). (C) 16S rRNA amplicon sequencing in the
642 colon of uncolonized (*Ctrl*) mice or mice colonized with *Tmu* or *Tc* for three weeks. The 15 most
643 abundant taxa are shown, and data is the average abundance from five mice per group. (D)
644 Relative abundance of bacteria depleted by *Tmu* and *Tc* colonization in each region of the
645 intestine, from 16S Sequencing. (E) Representative microscopy showing localization of *Tmu*

646 and *Tc* in the colons of mice fed complex chow. Scale bars are 100μm. (F) Quantification of
647 protist localization by microscopy. Higher mucus/lumen protist signal indicates tighter
648 localization to the colonic mucus layer. (G) Number of CAZymes encoded by protists and
649 bacterial competitors with predicted activity against host and plant glycans. (H) Abundance (in
650 TPM) of putative mucus-utilization genes in *Tc* and *Tmu* from transcriptomics data. CAZymes
651 were identified by dbCAN2. For putative mucolytic CAZymes the glycoside hydrolase family
652 (GH) is indicated in brackets. Graphs (A, B, D, F) depict mean, error bars depict SD in (D). Each
653 symbol represents an individual mouse from one to three pooled experiments (A). **p<0.01,
654 ***p<0.001, ****p<0.0001, NS, not significant by Mann-Whitney test (A, B) or Student's t test (F).
655
656 **Figure 6. Dietary MAC deprivation causes *Tmu* to switch to a mucolytic metabolism.**
657 (A) Representative images of *Tmu* in the colons of mice fed complex chow (left) and defined
658 chow with cellulose (def-C) chow (right) for two weeks. Scale bars are 100μm. (B) Quantification
659 of microscopy on *Tmu* localization to the mucus layer when mice are fed complex chow or def-C
660 chow. Higher mucus/lumen protist signal indicates tighter localization to the colonic mucus
661 layer. (C) Growth curves of *Tmu* cultured *in vitro*, using the newly developed PBF medium, with
662 maltose (cyan) or mucus (green) added as carbon sources, or with no defined carbon source
663 added (grey). (D) Differential gene expression from *Tmu* grown *in vitro* with either maltose or
664 mucus as a carbon source. Green data points indicate putative mucus-utilization genes with
665 increased expression when grown in the presence of mucus. GH families for putative mucolytic
666 CAZymes are labelled. Dashed lines delineate the fold change cutoff (1.5-fold). (E) *Tmu*
667 abundance in the stool of mice fed def-CI or def-C chow for two weeks, with or without an
668 antibiotic cocktail (VNA). Antibiotics used are Vancomycin, V; Neomycin, N; Ampicillin, A. (F)
669 *Tmu* abundance in the cecal contents of mice fed cellulose chow for two weeks with different
670 antibiotic treatments. (G) 16S rRNA sequencing of stool from *Tmu*-colonized mice fed def-C
671 chow for two weeks with different antibiotic treatments. Each column represents an individual

672 mouse. (H) Relative abundance of known mucolytic bacteria in mice from each group in (G). LD,
673 limit of detection. Graphs (B, C, E, F, H) are plotted as mean, error bars depict SD in (C). Each
674 symbol represents an individual mouse from one to three pooled experiments (B, E, and F).
675 **p<0.01, ****p<0.0001, NS, not significant by Mann-Whitney test (E, F) or Student's t test (B).
676

677 **Figure 7. Diet-protist interactions shape the host immune response.**

678 (A) *In vitro* growth curve of *Tmu* cultured in PBF medium with maltose or inulin added as
679 carbon sources, or with no defined carbon source added. (B) Representative images of *Tmu* in
680 the colons of mice fed complex chow or defined chow with cellulose and inulin (def-CI) for three
681 weeks. Scale bars are 50μm. (C) Quantification of microscopy on *Tmu* localization to the colonic
682 mucus layer when mice are fed complex chow or def-CI chow. Higher mucus/lumen protist
683 signal indicates tighter localization to the colonic mucus layer. (D) Relative abundance of
684 *Bacteroidetes* bacteria in the stool of mice colonized by *Tmu* or *Tc*, with the fermentable fiber
685 inulin or the non-fermentable fiber cellulose as the only fiber sources in the diet. (E) Maximum
686 *Tmu* culture density when grown *in vitro* with different dietary MACs added as carbon sources to
687 PBF medium. All statistical comparisons are relative to culture with no defined carbon source
688 added. (F) Differential expression of *Tmu* transcripts grown *in vitro* with either maltose or
689 arabinoxylan as a carbon source. Cyan data points represent upregulated CAZymes with
690 potential activity towards arabinoxylan. Dashed lines delineate the fold change cutoff (1.5-fold).
691 (G and H) Epithelial tuft cell (G) and lamina propria Th2 (H) frequency in the distal SI of mice
692 colonized with *Tmu* or no protist (*Ctrl*) and fed complex or def-CI chow for 3 weeks. (I) *Tmu*
693 abundance in the distal SI of mice fed complex chow and def-CI chow for 3 weeks. Data is
694 plotted as mean with SD (A and E). Each symbol represents an individual mouse from two or
695 three pooled experiments (C, D, G, H, and I). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, NS,
696 not significant by Mann-Whitney test (D, I) or Student's t test (C, E, G, H).
697

698 **STAR METHODS**

699 **Key Resources Table**

700 [See separate document]

701

702 **RESOURCE AVAILABILITY**

703 **Lead contact**

704 Further information and requests for resources and reagents may be directed to and will be

705 fulfilled by the lead contact, Michael Howitt (mhowitt@stanford.edu).

706 **Materials availability**

707 Mouse and microbial strains used in this study will be made available upon request addressed

708 to the lead contact.

709 **Data and code availability**

710 Raw sequencing files and genome assemblies have been deposited at Stanford Digital

711 Repository and are publicly available as of the date of publication (links to each dataset can be

712 found in Key Resources Table). Any additional information required to reanalyze the data

713 reported in this paper is available from the lead contact upon request.

714

715 **Animals**

716 C57BL/6 WT and *Caspase1*$^{-/-}$ mice were purchased from Jackson Laboratory. Mice from

717 Jackson Laboratory are protist-negative, as verified by PCR on the stool and examination of

718 cecal contents. *Tprm5*$^{-/-}$ and *Gpr91*$^{-/-}$ mice were generously provided by Dr. Robert Margolskee

719 and Amgen, respectively, under materials transfer agreements. The IL13 reporter mouse line

720 (SMART13 mice)[54] were generously provided by Dr. Jakob von Moltke. All mice were used at 6-

721 12 weeks of age. Experiments were performed using age and gender matched groups. All

722 animal procedures used in this study were approved by the Institutional Animal Care and Use

723 Committee (IACUC) at Stanford School of Medicine.

724 All mice were housed in individually ventilated cage systems and maintained at specific
725 pathogen free (SPF) health status, unless stated otherwise. Colonies of protist-negative and
726 protist-colonized mice were maintained and handled separately. Protist colonization status was
727 routinely verified by PCR on stool and examination of cecal contents by light microscopy. Cage
728 bottoms were covered with autoclaved bedding and enrichment material. Mice were handled
729 within HEPA filtered air exchange stations, and all surfaces and tools were sanitized using
730 Virkon™ S disinfectant before and after handling. All cage changes were performed by lab staff,
731 who were required to wear personal protection equipment (lab coats, gloves, disposable
732 sleeves) to handle mice used in this study.
733 Custom diets for *in vivo* fiber experiments were obtained from Research Diets, Inc. The
734 diets were the same as those used previously[21], with diet IDs D11112201 (Defined chow with
735 cellulose and inulin), D11112229 (no fiber), D17030102 (cellulose), D11112226 (inulin). For
736 antibiotic treatment experiments, mice were given antibiotics in the drinking water, which was
737 changed every 3 days. Antibiotics and concentrations used were: vancomycin (0.5g/mL),
738 neomycin (1g/mL), ampicillin (1g/mL). For succinate treatment, disodium succinate (100mM)
739 was added to the drinking water and changed every 3 days.
740
741 **Method details**
742 Identification of protists in mouse stool
743 Mice at Stanford University animal facilities were screened for commensal protists by extracting
744 DNA from stool using the DNeasy Powersoil Pro Kit (Qiagen) according to manufacturer
745 instructions. Samples were then used as input for PCR using pan-parabasalid primers (forward:
746 5'-CCACGGGTAGCAGCA-3', reverse: 5'-GGCAGGGACGTATTCAA-3'). These primers
747 generate an approximately 1.1kb amplicon of the parabasalid internally transcribed spacer (ITS)
748 sequence, which was sequenced by Sanger Sequencing for phylogenetic classification.
749 Phylograms were generated using MAFFT[55].

750 Isolation of *Tritrichomonas spp.* from mice

751 Protists were purified as described previously[19], with modifications. Briefly, cecal contents were
752 harvested from WT C57BL/6 mice and passed over a 40µm strainer. Contents were then
753 washed three times in PBS and purified at the interface of a 40%/80% Percoll (GE healthcare)
754 gradient, after centrifugation at 1000xg for 10minutes with no brake. The number of viable
755 protists was quantified using a hemocytometer, and $1x10^6$ trophozoites were inoculated per mL
756 of PBF growth medium. The PBF medium was supplemented with 1x Penicillin/Streptomycin
757 (Caisson Labs), 2.5µg/mL Amphotericin B (Sigma Aldrich), 100µg/mL Vancomycin (Alfa Aesar),
758 12.5µg/mL Chloramphenicol (Sigma). Protists were then cultured overnight in an anaerobic
759 chamber (Coy Labs) at 37˚C.

760 Colonization of mice with protists

761 Mice were colonized with commensal parabasalid protists as described previously[19] with slight
762 modifications. Briefly, protists were resuspended in PBS and mice were colonized with $1x10^5$
763 protists per mouse by oral feeding. Colonization was verified after 7 days by qPCR on the stool
764 using *Tmu* or *Tc* specific primers. For *Tc* colonization of mice, protists were isolated directly
765 from mouse ceca one day prior to colonization. For *Tmu* colonization of mice, protists were
766 cultured *in vitro* for no more than 20 passages prior to colonization of mice.

767 Scanning Electron Microscopy (SEM)

768 SEM was performed on *Tmu* and *Tc* as described previously[19]. Briefly, protists were isolated
769 from the ceca of SPF mice as described above and adhered to poly-lysine coated coverslips.
770 Samples were then fixed in 2.5% glutaraldehyde in 0.1M cacodylate pH 7.2. Samples were then
771 rinsed and post-fixed for 30min in 1% OsO_4 in 0.1M cacodylate buffer and dehydrated in a
772 graded series of ethanol prior to critical point drying with liquid CO_2. Protists were then sputter-
773 coated with 5nm platinum and examined with a Zeiss Sigma scanning electron microscope.

774 Collating previously published human gut metagenomic samples

775 Prevalence of parabasalid species across lifestyle was characterized using a curated collection
776 of 1,800 metagenomes including samples from “industrial”, “transitional” and “hunter-gatherer”
777 populations that were reported previously[28]. Due to the sensitivity of this analysis to read depth,
778 we removed studies that had a median read depth lower than 5 giga base pairs.

779 Discovery of putative novel ITS sequences

780 We compiled a set of Parabasalid internal transcribed spacer (ITS) sequences derived in this
781 study or from public databases. We used these sequences as “bait” to mine for novel ITS
782 sequences in the contigs of metagenome assemblies of samples derived from Tanzania, Nepal
783 or California[28]. The comparison between bait and assembly contigs was performed by
784 mashmap[56]. A phylogenetic tree of these putative novel ITS sequences, along with our bait ITS
785 sequences, was created using FastTree[57].

786 Species prevalence analysis

787 All reads were mapped against a database of Parabasalid ITS sequences derived in this study
788 or from public databases (Bowtie2)[58]. Resulting mappings were processed using inStrain profile
789 (v1.2.14)[59] and CoverM v0.4.0 (https://github.com/wwood/CoverM). Species with more than one
790 read aligning to the ITS sequence at a breadth greater than 20% were considered present and
791 prevalence was calculated as the percentage of metagenomes in which the species was
792 present.

793 High molecular weight DNA extraction and Nanopore Sequencing

794 To obtain pure protist DNA for genome sequencing, *Tc*-colonized mice were treated with an
795 antibiotic and antifungal cocktail in the drinking water (0.5g/L Vancomycin, 1g/L Neomycin, 1g/L
796 Ampicillin, 0.2g/L Amphotericin B) for 6 days. *Tc* was then purified from the cecal contents as
797 described above, with the modification of two rounds of Percoll purification, instead of one. For
798 *Tmu*, protists were instead axenically cultured. High molecular weight DNA from each protist
799 was then isolated by gently resuspending pelleted protists in lysis buffer (PBS+0.5% sodium
800 dodecyl sulfate), using a wide-bore pipet tip. Lysed cells were transferred to DNA LoBind tubes

801 (Eppendorf) and incubated with Monarch RNase A (NEB) for 2 hours at 37˚C. Next, Proteinase
802 K (Qiagen) was added, and samples were incubated for 2 hours at 50˚C. Nucleic acid was then
803 isolated using two rounds of extraction with Phenol:Chloroform:Isoamyl alcohol (Sigma). To
804 avoid pipetting during transfer steps, vacuum grease (Dow Corning) was added as a phase-lock
805 gel. Protein contaminants were then precipitated using 2M Ammonium Acetate, and then DNA
806 was precipitated overnight in ethanol. Next, DNA was pelleted by centrifugation at 8000xg, and
807 DNA pellets were washed three times with 70% ethanol. DNA was then resuspended in water
808 and used for library preparation and sequencing using the Nanopore Ligation Sequencing Kit
809 and Minion Flow Cell (Oxford Nanopore).

810 Genome assembly and annotation

811 Genome assemblies for both *T. casperi* and *T. musculis* were constructed using Flye[60], v. 2.9-
812 b1768 with default parameters. Short read libraries were then mapped against the genome
813 assemblies with Bowtie2[58], v. 2.4.4 and resulting alignments were used for assembly polishing
814 with pilon[61]. Genes were predicted with MetaEuk[62], v. 5.34c21f2 with Uniref90 as the reference
815 protein dataset. Genome completeness was assessed with BUSCO[63], v5.0.0 in protein mode
816 with Eukaryota ODB10 lineage marker gene set. Genome diagrams were generated using
817 DNAPlotter[64], v. 18.2.0.

818 RNA-Sequencing

819 RNA-Sequencing was performed on *Tmu* and *Tc in vivo* by first isolating the luminal contents
820 from the distal small intestine and cecum of 3 mice per group. Contents were strained over
821 40µm filters, then washed with ice-cold PBS to remove debris. To minimize potential for sample
822 processing to affect transcriptomes, samples were then immediately resuspended in Qiazol
823 (Qiagen). RNA was purified using the Direct-zol RNA Miniprep kit (Zymo Research). RNA-Seq
824 libraries were then generated using the KAPA Stranded RNA-Seq kit with poly-dT enrichment of
825 mRNA transcripts, and libraries were sequenced on an Illumina HiSeq instrument.

826 Transcriptomics on the *in vitro* grown *Tmu* samples was performed on protists grown for
827 3 days in the condition of interest. Cultures (3 biological replicates per condition) were pelleted
828 at 700xg for 5min at 4˚, then resuspended in Qiazol. RNA was then isolated using the Direct-zol
829 RNA Miniprep kit (Zymo Research), as above. Library preparation and sequencing was
830 performed at Azenta Life Sciences (South Plainfield, NJ). Libraries were constructed using the
831 TruSeq Stranded mRNA kit (Illumina), with poly-dT enrichment of mRNA transcripts. Samples
832 were sequenced on a HiSeq instrument (Illumina).
833 Analysis of all transcriptomics datasets was performed using alignment-based analysis,
834 using the newly sequenced *Tmu* and *Tc* genomes. Briefly, Illumina adapters were trimmed and
835 low-quality sequence was removed using Trimmomatic[65]. Reads were then aligned to the
836 respective genome using Burrows Wheeler Aligner[66]. Samtools[67] was used to convert
837 corresponding files to binary alignment and map (bam) files and sort the reads by position. Next,
838 featureCounts[68] was used to quantify read depth of each genome feature. Differential
839 expression analysis was performed using DESeq2[69].
840 Protist enumeration
841 Enumeration of protists from the small intestinal tissue was performed as described
842 previously[19], with modifications. Briefly, the distal 10cm of small intestine was removed and
843 flushed with ice-cold sterile PBS using a 19-gauge feeding needle. The contents were then
844 pelleted and the supernatant was aspirated. Genomic DNA was isolated using the Powersoil
845 Pro Kit (Qiagen) according to manufacturer instructions. To detect and enumerate protists,
846 quantitative PCR (qPCR) was performed using PowerUp SYBR Green Master Mix (Applied
847 Biosystems). For *T. musculis*, the following primers specific to the 28S rRNA gene were used:
848 5'-GCTTTTGCAAGCTAGGTCCC-3'and 5'- TTTCTGATGGGGCGTACCAC-3'[19]. For *T. casperi*
849 detection, primers were designed that specifically recognized this protist: 5'-
850 AGGTTACTGAATCATACATGCGT-3' and 5'-GCAGGAGTTGCTTTCATTGTG-3'. qPCRs were
851 run on a QuantStudio 3 Real Time PCR instrument (Thermo Fisher). To convert qPCR values

852 into protist numbers, protists were isolated and counted using a hemocytometer before
853 extracting genomic DNA and analyzing a dilution series by qPCR. These results were used to
854 create a standard curve, and regression analysis was used to convert Ct values to protist
855 numbers, as has been described previously[15,19].

856 Quantification of protists in the stool was performed similarly. Stool pellets were
857 collected from mice and weighed. A fraction of the stool was dried and used to calculate the
858 water content of the stool, while the rest was used for DNA extraction. Quantification was
859 performed using the same qPCR primers as above. Protists were quantified in the cecum by
860 weighing the cecal luminal contents, then resuspending in PBS. Live trophozoites were then
861 counted using a hemocytometer and normalized to cecal weight.

862 Isolation of Epithelial Cells for Flow Cytometry

863 Ileal epithelial cells were isolated as described previously[19]. Briefly, the terminal 10cm of the
864 small intestine was removed and trimmed of fat. Luminal contents were gently flushed with
865 10mL of ice-cold PBS, then the tissue was opened longitudinally. Peyer's Patches were
866 removed and the tissue was incubated on ice in PBS, 5mM HEPES pH 7.4, 2% heat-inactivated
867 fetal bovine serum (iFBS), 1mM DTT. Tissue was then transferred to pre-warmed PBS, 5mM
868 HEPES, 2% iFBS, 5mM EDTA and shaken at 200rpm for 15 minutes at 37˚, followed by
869 vigorous shaking. This was repeated and epithelial cells from both fractions were combined and
870 washed with PBS. Tissue was then digested in DMEM containing 10% iFBS, 0.5U/mL Dispase
871 II (StemCell Technologies), 50µg/mL DNase (Roche) for 10 minutes at 37˚C. The cells were
872 then filtered over 40µm strainers and washed with PBS, 2% iFBS, 1mM EDTA. Cells were then
873 incubated for 10 minutes with anti-CD16/CD32 (clone 93, Biolegend) and then stained with the
874 following antibodies: PacBlue-conjugated anti-CD45 (clone 30-F11, Biolegend), PE/Cy7-
875 conjugated anti-EpCam (clone G8.8, Biolegend), and Alex647-conjugated anti-SiglecF (clone
876 E50-2440, BD Biosciences). Cell viability was assessed by propidium iodide staining
877 (Biolegend). Cells were then analyzed on a BD FACSCanto II.

878 Lamina propria cell isolation and flow cytometry

879 The distal 10cm of small intestine, or the entire colon, were collected and epithelial cells were
880 removed as described above. The remaining tissue was then minced and digested in RPMI
881 10% iFBS, 0.1 units/mL Dispase II (StemCell Technologies), 50µg/mL DNase (Roche),
882 collagenase A (Roche) (0.25mg/mL for ilea, 0.5mg/mL for colons). Tissue was digested for 30
883 minutes followed by a second digestion for 40 minutes. Digested tissue was passed through a
884 40µm strainer, and the cells were counted using a Cellometer Auto T4 (Nexcelom Bioscience).
885 For cytokine measurements, cells were stimulated in RPMI supplemented with 10% iFBS,
886 Pen/Strep, PMA, Ionomycin, and GolgiPlug. Cells were incubated at 37˚ for 4 hours to allow
887 stimulation to occur before being harvested. Cells were then incubated with anti-CD16/32 (clone
888 92, BioLegend) before surface staining. For T cell analysis, cells were stained with PacBlue-
889 conjugated anti-CD45 (clone 30-F11, BioLegend), FITC-conjugated anti-CD4 (clone GK1.5,
890 BioLegend), and PerCP/Cy5.5-conjugated anti-CD3 (clone 17A2, BioLegend). For ILC2
891 analysis, cells were stained with PacBlue-conjugated anti-CD45 (clone 30-F11, BioLegend),
892 PerCP/Cy5.5-conjugated anti-CD127 (clone A7R34, BioLegend), PE/Cy7-conjugated anti-
893 KLRG1 (clone 2F1, BioLegend), PE-conjugated anti-human CD4/Smart13 (clone RPA-T4,
894 BioLegend), and a custom FITC conjugated lineage cocktail of antibodies against NK1.1 (clone
895 PK136, BioLegend), CD19 (clone 6D5, BioLegend), CD4 (clone GK1.5, BioLegend), CD8 (clone
896 53-6.7, BioLegend), CD64 (clone S18017D, BioLegend), CD11b (clone M1/70, BioLegend), and
897 TER119 (clone TER-119, BioLegend). Cells were then stained with the viability dye Zombie-NIR
898 (BioLegend) according to manufacturer instructions prior to fixation with the Foxp3 Transcription
899 factor staining buffer set (eBioscience). For T cell analysis, fixed cells were permeabilized and
900 stained with PE-conjugated anti-IL17 (clone 9B10, Biolegend) and APC-conjugated anti-IFNγ
901 (clone XMG1.2, eBioscience), or with Alexa647-conjugated anti-GATA3 (clone L50 823, BD) or
902 Alexa647-conjugated anti-FOXP3 (clone MF-14, BioLegend). Cells were then analyzed on a BD
903 FACSCanto II.

904 Tissue preparation and staining for immunofluorescence

905 Tissue sections of interest were removed and fixed in methacarn fixative (Fisher Scientific) for
906 48 hours. Tissue was then washed three times in 100% ethanol, followed by three washes in
907 Histo-Clear II (Electron Microscopy Sciences). Fixed tissue was embedded in paraffin wax prior
908 to sectioning (4μm) using a microtome. For immunofluorescence, tissue was first deparaffinized
909 as has been described previously[70], then blocked with blocking buffer (PBS supplemented with
910 3% BSA, 0.1% Saponin, 2% goat serum). Sections were then stained with the anti-*Tmu*
911 antibody (generated by Rockland Scientific) and anti-E Cadherin (clone 36/E) overnight. After 3
912 washes in PBS, sections were stained with secondary antibodies fluorescently conjugated to
913 Alexa488 (*Tmu*) and Alexa660 (E Cadherin). DAPI (Sigma Aldrich) and rhodamine-conjugated
914 Ulex Europaeus Agglutinin I (Vector Laboratories) were added during the secondary staining
915 step as well. After 3 PBS washes, tissue was mounted with Vectashield (Vector Labs) and
916 imaged on a Zeiss LSM 700 confocal microscope.

917 Microscopy quantification of protist localization to the colonic mucus layer

918 One to three images were taken for each mouse for quantification. A 40μm tall mucosal region
919 of interest (ROI) was drawn at the luminal edge of the inner colonic mucus layer. An ROI of
920 identical size was placed in a representative region of the lumen, a minimum of 100μm away
921 from the mucus layer. Fluorescence signal from the protist channel was measured, and the ratio
922 of the signal from the mucosal ROI to the luminal ROI was calculated. The average of the ratios
923 from each image per mouse was then calculated as the mucus/lumen signal.

924 Metabolite measurements

925 For intracellular metabolite analysis, mice were treated with an antibiotic cocktail in the drinking
926 water (0.5g/L Vancomycin, 1g/L Neomycin, 1g/L Ampicillin, 0.2g/L Amphotericin B) for 6 days.
927 Protists were then isolated as described above and lysed in 1mL 80% methanol. Cellular debris
928 was removed through centrifugation, followed by filtration using 0.22μm filters, snap frozen in
929 liquid nitrogen, and stored at -80˚. Prior to liquid chromatography mass spectrometry (LC/MS)

930 analysis, intracellular extracts (800 µL) were dried in a vacuum concentrator at room
931 temperature for approximately 4 hours then resuspended in 50 µL of HPLC-grade water spiked
932 with ^{13}C-labeled metabolites as an internal standard.
933 For metabolite measurements from cecal supernatants, cecal contents were squeezed
934 into tubes and weighed, prior to gentle resuspension in sterile saline solution (0.15M NaCl).
935 Samples were then pelleted at 500xg for 5min at 4˚, and supernatants were transferred to clean
936 tubes. Samples were then pelleted at 2000xg for 10min at 4˚, and then supernatants were
937 filtered over 0.22µm filters, snap frozen in liquid nitrogen, and stored at -80˚. Prior to LC/MS
938 analysis, cecal extracts were diluted in HPLC-grade water spiked with ^{13}C-labeled metabolites
939 as an internal standard.
940 For organic acid measurements, samples (5 µL) were run in reverse phase on a 2.1 mm
941 x 150 mm Waters Acquity UPLC BEH C18 column with 1.7 µm packing. The column was fitted
942 with a Vanguard pre-column of the same composition. Analytes were separated on an Agilent
943 1290 Infinity II UPLC (binary pumps) and detected using an Agilent 6545 LC/MS Quadrupole
944 Time-of-Flight (qTOF) instrument equipped with a dual jet stream electrospray ionization source
945 (ESI) operating under extended dynamic range (EDR 1700 m/z) in the negative (ESI-) ionization
946 mode. The mobile phase solvents consisted of 20 mM ammonium formate pH 2.9 (A) and
947 acetonitrile (B). The separation began with 99.5% A and 0.5% B for 4 min at 0.4 mL/min, then B
948 was increased to 99.5% at 0.5 mL/min and held for 1.5 min, at which time the flow rate was
949 increased to 0.8 mL/min. At 11.5 min, solvent B was reduced to 0.5% and the flow rate was
950 reduced to 0.5 mL/min for 2.5 min, at which time it was further reduced to 0.4 mL/min. The
951 separation ended at 14 min. Specific ion source parameters included: Drying gas temperature
952 was 150˚C at 12L/min, sheath gas temperature was 325˚C at 12 L/min, Nebulizer was at 45
953 psig, capillary voltage was 2000 V, nozzle voltage was 0V, fragmentor voltage was 90 V,
954 skimmer voltage was at 35 V, octopole RF was 750V. For quantitation purposes, standard
955 curves of all analytes were prepared in HPLC water spiked with ^{13}C-labeled metabolites as an

956 internal standard. Metabolite levels were determined by manually integrating the area under
957 each chromatographic peak and their concentrations in the samples determined by interpolating
958 the peak area for each compound on the corresponding standard curve.

959 For short chain fatty acid measurements, a LC-MS-based SCFA/OA quantification
960 method was adapted[71]. Briefly, cell pellets diluted in extraction buffer containing: 80% HPLC-
961 grade water (Fisher), 20% HPLC-grade acetonitrile (ACN; Fisher), and labeled isotopes of each
962 SCFA/OA measured (2.5 μM d3-acetic acid [Sigma-Aldrich], 1 μM propionic-3,3,3-d3 acid [CDN
963 Isotopes], and 0.5 μM butyric-4,4,4-d3 acid [CDN Isotopes]). Acid-washed beads were added to
964 the samples and the samples were subsequently shaken at 30 Hz for 10 minutes. The
965 supernatants were then derivatized in a combination of 200 mM 3-nitrophenylhydrazine
966 hydrochloride (Sigma-Aldrich; dissolved in 50% ACN and 50% water) and 120 mM 1-ethyl-3-(3-
967 dimethylaminopropyl)carbodiimide hydrochloride (Pierce; dissolved in 47% ACN, 47% water,
968 and 6% HPLC-grade pyridine [Sigma-Aldrich]). The reaction was incubated at 37°C for 30 min
969 and subsequently diluted 50X prior to mass spectrometry analysis.

970 Analyses were carried out using an Agilent 6470 triple-quadrupole LC-MS and a Waters
971 Acquity ultraperformance liquid chromatography (UPLC) ethylene-bridged hybrid (BEH) C18
972 column (100-mm length, 2.1-mm inner diameter, 130-Å pore size, 1.7-μm particle size). The
973 mobile phases were water-formic acid (100:0.01, vol/vol; solvent A) and acetonitrile-formic acid
974 (100:0.01, vol/vol; solvent B). Quantification of analytes was done by standard isotope dilution
975 protocols. In brief, serial dilutions of a SCFA/OA standard solution (10 mM, 1 mM, 0.1 mM,
976 0.01 mM, 0.001 mM, and 0 mM) were derivatized as described above and included in each run
977 to verify that sample concentrations were within linear ranges. For samples within linear range,
978 analyte concentration was calculated as the product of the paired internal standard
979 concentration and the ratio of analyte peak area to internal standard peak area.

980 Untargeted LC-MS metabolomics was performed as described previously[36] using $C18^+$
981 and $C18^-$ modes.

982 Dietary fiber experiments

983 For experiments involving manipulation of dietary fiber, mice (fed normal complex chow) were
984 colonized by $1x10^5$ protists between 4-6 weeks of age. Ten days later, the diets were changed
985 to chows with defined fiber compositions. Stool was collected for protist quantification two
986 weeks later, and mice were euthanized for collection of tissue for additional analyses. For
987 experiments involving manipulation of dietary fiber and antibiotic treatment, mice were colonized
988 by protists as above. Ten days later, mice were started on the indicated antibiotic treatment,
989 followed by diet switch three days after the start of antibiotics.

990 16S Bacterial Profiling

991 For 16S bacterial profiling, mice were colonized by $1x10^5$ protists between 4-6 weeks of age,
992 and stool was collected 3 weeks later for microbiome analysis. DNA was extracted using the
993 Powersoil Pro Kit (Qiagen). The DNA samples were prepared for targeted sequencing at Zymo
994 Research (Irvine, CA). Libraries were made using the *Quick*-16S NGS Library Prep Kit using
995 Primer Set V3-V4 (Zymo Research, Irvine, CA). Samples were sequenced on an Illumina MiSeq
996 with a v3 reagent kit. Unique amplicon sequences were inferred from raw reads using the
997 Dada2 pipeline[72]. Chimeric sequences were removed, and taxonomy assignment was
998 performed using Uclust from Qiime2[73]. Taxonomy was assigned with the Zymo Research 16S
999 Database. Alpha-diversity and composition analyses were performed with Qiime.

1000 Absolute abundance quantification was performed using qPCR. A standard curve was
1001 made with plasmid DNA containing one copy of the 16S gene, prepared in 10-fold serial
1002 dilutions. The primers used were the same as those used in targeted library preparation. The
1003 equation generated by the plasmid DNA standard curve was used to calculate the number of
1004 gene copies in the reaction for each sample. The PCR input volume was used to calculate the
1005 number of gene copies in each DNA sample. The number of genome copies per microliter of
1006 DNA in each sample was calculated by dividing the gene copy number by an assumed number
1007 of gene copies per genome. The value used for 16S copies is 4. The amount of DNA per

1008 microliter of sample was calculated using an assumed genome size of 4.64×10^6bp, the genome
1009 size of *Escherichia coli*. The mass of the stool sample used for DNA extraction was then used to
1010 calculate the number of genome copies per milligram of stool.

1011 Bacterial relative abundance measurements by qPCR

1012 *Akkermansia muciniphila* and *Bacteroidetes* quantification was performed using DNA isolated
1013 from stool as described above. Relative abundance of bacterial groups were then quantified
1014 using the following primers: *Akkermansia muciniphila* 5'-CAGCACGTGAAGGTGGGGAC-3', 5'-
1015 CCTTGCGGTTGGCTTCAGAT-3'[74]. *Bacteroidetes* 5'-GTTTAATTCGATGATACGCGAG-3', 5'-
1016 TTAASCCGACACCTCACGG-3'[75]. Universal 16S 5'-AAACTCAAAKGAATTGACGG-3', 5'-
1017 CTCACRRCACGAGCTGAC-3'[75]. qPCRs were performed using PowerUp SYBR Green Master
1018 Mix (Applied Biosystems) on a QuantStudio 3 Real-Time PCR machine (Thermo Fisher). Ct
1019 values for bacterial groups of interest were then normalized to total 16S for relative abundance
1020 analysis.

1021 Quantification and Statistical Analysis

1022 Groups were compared using Prism v9 software (GraphPad) using the two-tailed unpaired
1023 Student's t test or Mann-Whitney test where appropriate. Differences were considered
1024 statistically significant when $p \leq 0.05$.

1025

1026

1027

1028

1029

1030

1031

1032

1033

1034 **References**

1035 1. Chu, H. & Mazmanian, S. K. Innate immune recognition of the microbiota promotes host-
1036 microbial symbiosis. *Nat. Immunol.* **14**, 668–675 (2013).
1037 2. Hooper, L. V., Littman, D. R. & Macpherson, A. J. Interactions between the microbiota and
1038 the immune system. *Science* **336**, 1268–1273 (2012).
1039 3. Kamada, N., Seo, S. U., Chen, G. Y. & Núñez, G. Role of the gut microbiota in immunity
1040 and inflammatory disease. *Nat. Rev. Immunol.* **13**, 321–335 (2013).
1041 4. Littman, D. R. & Pamer, E. G. Role of the Commensal Microbiota in Normal and Pathogenic
1042 Host Immune Responses. *Cell Host Microbe* **10**, 311–323 (2011).
1043 5. Rooks, M. G. & Garrett, W. S. Gut microbiota, metabolites and host immunity. *Nat. Publ.*
1044 *Group* **16**, 341–352 (2016).
1045 6. Atarashi, K. *et al.* Treg induction by a rationally selected mixture of Clostridia strains from
1046 the human microbiota. *Nature* **500**, 232–236 (2013).
1047 7. Belkaid, Y. & Hand, T. W. Role of the Microbiota in Immunity and Inflammation. *Cell* **157**,
1048 121–141 (2014).
1049 8. Levy, M. *et al.* Microbiota-Modulated Metabolites Shape the Intestinal Microenvironment by
1050 Regulating NLRP6 Inflammasome Signaling. *Cell* **163**, 1428–1443 (2015).
1051 9. Smith, P. M. *et al.* The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic
1052 Treg Cell Homeostasis. *Science* **341**, 569–573 (2013).
1053 10. Zheng, Y. *et al.* Bile acid metabolites control TH17 and Treg cell differentiation. *Nature* **576**,
1054 143–148 (2019).
1055 11. Bae, M. *et al.* Akkermansia muciniphila phospholipid induces homeostatic immune
1056 responses. *Nature* 1–6 (2022) doi:10.1038/s41586-022-04985-7.
1057 12. Ivanov, I. I. *et al.* Induction of Intestinal Th17 Cells by Segmented Filamentous Bacteria. *Cell*
1058 **139**, 485–498 (2009).

1059 13. Mazmanian, S. K., Round, J. L. & Kasper, D. L. A microbial symbiosis factor prevents
1060 intestinal inflammatory disease. *Nature* **453**, 620–625 (2008).
1061 14. Billy, V. *et al.* Blastocystis Colonization Alters the Gut Microbiome and, in Some Cases,
1062 Promotes Faster Recovery From Induced Colitis. *Front. Microbiol.* **12**, 641483 (2021).
1063 15. Chudnovskiy, A. *et al.* Host-Protozoan Interactions Protect from Mucosal Infections through
1064 Activation of the Inflammasome. *Cell* **167**, 444-456.e14 (2016).
1065 16. Clemente, J. C., Ursell, L. K., Parfrey, L. W. & Knight, R. The Impact of the Gut Microbiota
1066 on Human Health: An Integrative View. *Cell* **148**, 1258–1270 (2012).
1067 17. Lukeš, J., Stensvold, C. R., ková, K. J.-P. i & Parfrey, L. W. Are Human Intestinal
1068 Eukaryotes Beneficial or Commensals? **11**, e1005039 (2015).
1069 18. Parfrey, L. W. *et al.* Communities of microbial eukaryotes in the mammalian gut within the
1070 context of environmental eukaryotic diversity. *Front. Microbiol.* **5**, 429 (2014).
1071 19. Howitt, M. R. *et al.* Tuft cells, taste-chemosensory cells, orchestrate parasite type 2
1072 immunity in the gut. *Science* **351**, 1329–1333 (2016).
1073 20. Nadjsombati, M. S. *et al.* Detection of Succinate by Intestinal Tuft Cells Triggers a Type 2
1074 Innate Immune Circuit. *Immunity* **49**, 33-41.e7 (2018).
1075 21. Schneider, C. *et al.* A Metabolite-Triggered Tuft Cell-ILC2 Circuit Drives Small Intestinal
1076 Remodeling. *Cell* **174**, 271-284.e14 (2018).
1077 22. Escalante, N. K. *et al.* The common mouse protozoa Tritrichomonas muris alters mucosal T
1078 cell homeostasis and colitis susceptibility. *J. Exp. Med.* **213**, 2841–2850 (2016).
1079 23. Brune, A. Symbiotic digestion of lignocellulose in termite guts. *Nat. Rev. Microbiol.* **12**, 168–
1080 180 (2014).
1081 24. Muller, M. *et al.* Biochemistry and Evolution of Anaerobic Energy Metabolism in Eukaryotes.
1082 *Microbiol. Mol. Biol. Rev.* **76**, 444–495 (2012).
1083 25. Petrin, D., Delgaty, K., Bhatt, R. & Garber, G. Clinical and Microbiological Aspects
1084 ofTrichomonas vaginalis. *Clin. Microbiol. Rev.* **11**, 300–317 (1998).

1085 26. Cepicka, I., Hampl, V. & Kulda, J. Critical Taxonomic Revision of Parabasalids with
1086 Description of one New Genus and three New Species. *Protist* **161**, 400–433 (2010).
1087 27. Smits, S. A. *et al.* Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of
1088 Tanzania. *Science* **357**, 802–806 (2017).
1089 28. Carter, M. M. *et al.* Ultra-deep sequencing of Hadza hunter-gatherers recovers vanishing
1090 gut microbes. *Cell* **186**, 3111-3124.e13 (2023).
1091 29. Atarashi, K. *et al.* Th17 Cell Induction by Adhesion of Microbes to Intestinal Epithelial Cells.
1092 *Cell* **163**, 367–380 (2015).
1093 30. Westrop, G. D. *et al.* Metabolomic profiling and stable isotope labelling of Trichomonas
1094 vaginalis and Tritrichomonas foetus reveal major differences in amino acid metabolism
1095 including the production of 2-hydroxyisocaproic acid, cystathionine and S-methylcysteine.
1096 *PLOS ONE* **12**, e0189072 (2017).
1097 31. Barratt, J., Gough, R., Stark, D. & Ellis, J. Bulky Trichomonad Genomes: Encoding a Swiss
1098 Army Knife. *Trends Parasitol.* **32**, 783–797 (2016).
1099 32. Carlton, J. M. *et al.* Draft genome sequence of the sexually transmitted pathogen
1100 Trichomonas vaginalis. *Science* **315**, 207–212 (2007).
1101 33. Palmieri, N., Ramires, M. de J., Hess, M. & Bilic, I. Complete genomes of the eukaryotic
1102 poultry parasite Histomonas meleagridis: linking sequence analysis with virulence /
1103 attenuation. *BMC Genomics* **22**, 753 (2021).
1104 34. Ferreyra, J. A. *et al.* Gut microbiota-produced succinate promotes C. Difficile infection after
1105 antibiotic treatment or motility disturbance. *Cell Host Microbe* **16**, 770–777 (2014).
1106 35. Fischbach, M. A. & Sonnenburg, J. L. Eating for two: How metabolism establishes
1107 interspecies interactions in the gut. *Cell Host Microbe* **10**, 336–347 (2011).
1108 36. Han, S. *et al.* A metabolomics pipeline for the mechanistic interrogation of the gut
1109 microbiome. *Nature* **595**, 415–420 (2021).

1110 37. Coyte, K. Z. & Rakoff-Nahoum, S. Understanding Competition and Cooperation within the
1111 Mammalian Gut Microbiome. *Curr. Biol.* **29**, R538–R544 (2019).
1112 38. Park, S.-Y. *et al.* Strain-level fitness in the gut microbiome is an emergent property of
1113 glycans and a single metabolite. *Cell* **185**, 513-529.e21 (2022).
1114 39. Tuncil, Y. E. *et al.* Reciprocal Prioritization to Dietary Glycans by Gut Bacteria in a
1115 Competitive Environment Promotes Stable Coexistence. *mBio* **8**, e01068-17 (2017).
1116 40. Porter, Nathan T & Martens, Eric C. The Critical Roles of Polysaccharides in Gut Microbial
1117 Ecology and Physiology. (2017).
1118 41. Sonnenburg, E. D. & Sonnenburg, J. L. Starving our Microbial Self: The Deleterious
1119 Consequences of a Diet Deficient in Microbiota-Accessible Carbohydrates. *Cell Metab.* **20**,
1120 779–786 (2014).
1121 42. Wardman, J. F., Bains, R. K., Rahfeld, P. & Withers, S. G. Carbohydrate-active enzymes
1122 (CAZymes) in the gut microbiome. *Nat. Rev. Microbiol.* 1–15 (2022) doi:10.1038/s41579-
1123 022-00712-1.
1124 43. Desai, M. S. *et al.* A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus
1125 Barrier and Enhances Pathogen Susceptibility. *Cell* **167**, 1339-1353.e21 (2016).
1126 44. Tailford, L. E., Crost, E. H., Kavanaugh, D. & Juge, N. Mucin glycan foraging in the human
1127 gut microbiome. *Front. Genet.* **6**, 81 (2015).
1128 45. Coyte, K. Z., Schluter, J. & Foster, K. R. The ecology of the microbiome: Networks,
1129 competition, and stability. *Science* **350**, 663–666 (2015).
1130 46. Patnode, M. L. *et al.* Interspecies Competition Impacts Targeted Manipulation of Human Gut
1131 Bacteria by Fiber-Derived Glycans. *Cell* **179**, 59-73.e13 (2019).
1132 47. Wu, G. *et al.* Gastrointestinal Tract and Dietary Fiber Driven Alterations of Gut Microbiota
1133 and Metabolites in Durco × Bamei Crossbred Pigs. *Front. Nutr.* **8**, 806646 (2022).
1134 48. Milani, C. *et al.* Bifidobacteria exhibit social behavior through carbohydrate resource sharing
1135 in the gut. *Sci. Rep.* **5**, 15782 (2015).

1136 49. Zhang, H. *et al.* dbCAN2: a meta server for automated carbohydrate-active enzyme
1137 annotation. *Nucleic Acids Res.* **46**, gky418 (2018).
1138 50. Holden, H. M., Rayment, I. & Thoden, J. B. Structure and Function of Enzymes of the Leloir
1139 Pathway for Galactose Metabolism*. *J. Biol. Chem.* **278**, 43885–43888 (2003).
1140 51. Sonnenburg, E. D. *et al.* Specificity of Polysaccharide Use in Intestinal Bacteroides Species
1141 Determines Diet-Induced Microbiota Alterations. *Cell* **141**, 1241–1252 (2010).
1142 52. Bellali, S. *et al.* Running after ghosts: are dead bacteria the dark matter of the human gut
1143 microbiota? *Gut Microbes* **13**, 1–12 (2021).
1144 53. Nayfach, S., Shi, Z. J., Seshadri, R., Pollard, K. S. & Kyrpides, N. C. New insights from
1145 uncultivated genomes of the global human gut microbiome. *Nature* **568**, 505–510 (2019).
1146 54. Liang, H.-E. *et al.* Divergent expression patterns of IL-4 and IL-13 define unique functions in
1147 allergic immunity. *Nat. Immunol.* **13**, 58–66 (2012).
1148 55. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence
1149 alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166
1150 (2019).
1151 56. Jain, C., Koren, S., Dilthey, A., Phillippy, A. M. & Aluru, S. A fast adaptive algorithm for
1152 computing whole-genome homology maps. *Bioinformatics* **34**, i748–i756 (2018).
1153 57. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 – Approximately Maximum-Likelihood
1154 Trees for Large Alignments. *PLoS ONE* **5**, e9490 (2010).
1155 58. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**,
1156 357–359 (2012).
1157 59. Olm, M. R. *et al.* inStrain profiles population microdiversity from metagenomic data and
1158 sensitively detects shared microbial strains. *Nat. Biotechnol.* **39**, 727–736 (2021).
1159 60. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads
1160 using repeat graphs. *Nat. Biotechnol.* **37**, 540–546 (2019).

1161 61. Walker, B. J. *et al.* Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection
1162 and Genome Assembly Improvement. *PLoS ONE* **9**, e112963 (2014).
1163 62. Karin, E. L., Mirdita, M. & Söding, J. MetaEuk—sensitive, high-throughput gene discovery,
1164 and annotation for large-scale eukaryotic metagenomics. *Microbiome* **8**, 48 (2020).
1165 63. Manni, M., Berkeley, M. R., Seppey, M. & Zdobnov, E. M. BUSCO: Assessing Genomic
1166 Data Quality and Beyond. *Curr. Protoc.* **1**, e323 (2021).
1167 64. Carver, T., Thomson, N., Bleasby, A., Berriman, M. & Parkhill, J. DNAPlotter: circular and
1168 linear interactive genome visualization. *Bioinformatics* **25**, 119–120 (2009).
1169 65. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence
1170 data. *Bioinformatics* **30**, 2114–2120 (2014).
1171 66. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform.
1172 *Bioinformatics* **25**, 1754–1760 (2009).
1173 67. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008
1174 (2021).
1175 68. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for
1176 assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
1177 69. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
1178 RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
1179 70. Earle, K. A. *et al.* Quantitative Imaging of Gut Microbiota Spatial Organization. *Cell Host*
1180 *Microbe* **18**, 478–488 (2015).
1181 71. Pensinger, D. A. *et al.* Butyrate Differentiates Permissiveness to Clostridioides difficile
1182 Infection and Influences Growth of Diverse C. difficile Isolates. *Infect. Immun.* **91**, e00570-22
1183 (2023).
1184 72. Callahan, B. J. *et al.* DADA2: High resolution sample inference from Illumina amplicon data.
1185 *Nat. Methods* **13**, 581–583 (2016).

1186 73. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science
1187 using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
1188 74. Schneeberger, M. *et al.* Akkermansia muciniphila inversely correlates with the onset of
1189 inflammation, altered adipose tissue metabolism and metabolic disorders during obesity in
1190 mice. *Sci. Rep.* **5**, 16643 (2015).
1191 75. Yang, Y.-W. *et al.* Use of 16S rRNA Gene-Targeted Group-Specific Primers for Real-Time
1192 PCR Analysis of Predominant Bacteria in Mouse Feces. *Appl. Environ. Microbiol.* **81**, 6749–
1193 6756 (2015).
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211

1212 **Supplemental figure legends**

1213 **Figure S1. Discovery of novel human- and mouse-associated parabasalids, related to**
1214 **Figure 1**.

1215 (A) Length of *Tmu* and *Tc* trophozoites based on SEM images. (B) Alignment of partial ITS
1216 sequences from *Tmu* and *Tc*. (C) *Tc* abundance in the stool of mice colonized for 3 weeks or 3
1217 months. Data is plotted as mean and each symbol represents an individual mouse.
1218 ****$p<0.0001$, NS, not significant by Students T test. (D) Average Nucleotide Identity (ANI) of
1219 the parabasalid reads in human metagenomic samples from (E). Imperfect ANI to the specified
1220 reference sequence indicates a relative of the reference protist. First author of the cohort study,
1221 country of origin, and population type ("Industrial", "Transitional", Hunter-Gatherer ("H-G") is
1222 indicated. Columns and rows are hierarchically clustered based on Euclidean distance. Cells
1223 corresponding to protists not identified in a population are colored grey. (E) Prevalence of
1224 parabasalids in human metagenomics datasets using mapping-based analysis, without
1225 exclusion criteria based on sequencing depth. Parabasalids with imperfect matches in (D)
1226 represent relatives of the reference protist species. (F) Percentage prevalence of *D. fragilis* as a
1227 function of sequencing depth, separated by industrialization status. (Abbreviations:
1228 TZA=Tanzania, NEP=Nepal, MNG=Mongolia, USA=United States of America,
1229 MDG=Madagascar, KAZ=Kazakhstan, SWE=Sweden, HMP=Human Microbiome Project,
1230 IND=Indonesia, ISR=Israel, FJI=Fiji, ETH=Ethiopia, THA=Thailand, PER=Peru,
1231 CMR=Cameroon).

1232

1233 **Figure S2. *Tmu* dominates tuft cell responses in the distal SI, related to Figure 2**

1234 A) Gating strategy for the identification of tuft cells in the epithelium of the distal SI by flow
1235 cytometry (Live $CD45^-$ $EpCAM^+$ $SiglecF^+$). (B) Gating strategy for the identification of ILC2s in
1236 the lamina propria (LP) of the distal SI by flow cytometry in SMART13 mice (Live $CD45^+$ $CD127^+$
1237 $KLRG1^+$ $IL\text{-}13^+$). (C) Gating strategy for the identification of Th1, Th2, and Th17 cells in the LP

1238 by flow cytometry: Live $CD45^+$ $CD3^+$ $CD4^+$ cells that are IFNγ $^+$ (Th1), $GATA3^+$ (Th2), or IL-17^+
1239 (Th17). (D) Protist quantification in each region of the intestine by qPCR, in SPF mice singly
1240 colonized with each protist. (E) Tuft cell frequency in the epithelium of mice colonized by each
1241 protist individually or colonized by both species (Co-col). (F) Protist quantification from *Tmu* and
1242 *Tc* co-colonized mice in each region of the intestine. Data points are average of at least 5 mice
1243 per group (D, F). Error bars show standard deviation (D, F). *p<0.05, **p<0.01, ***p<0.001, NS,
1244 not significant by Mann-Whitney test (D, F) or Student's t test (E).
1245
1246 **Figure S3: *Tmu* and *Tc* are sufficient to stimulate Th1 and Th17 immunity through an**
1247 **adhesion-independent manner, related to Figure 3.**
1248 (A) Frequency (top) and absolute abundance (bottom) of IFNγ, IL17, and IFNy-IL17 double
1249 positive cells in the colonic lamina propria of gnotobiotic mice monocolonized with each protist
1250 for three weeks, or left germ free (GF). (B) Abundance of regulatory T cells in the colonic lamina
1251 propria of wild type SPF mice colonized with each protist, or uncolonized mice (*Ctrl*). (C)
1252 Western Blot of *Tmu* and *Tc* lysates stained with the anti-*Tmu* polyclonal antibody. (D)
1253 Representative image of a *Tmu* trophozoite (left) and pseudocyst (right). Protists stained with
1254 anti-*Tmu* antibody (cyan), nuclei stained with DAPI (blue). Scale bars are 5μm. (E)
1255 Representative image of a *Tmu* cell adhered to the colonic epithelium. Scale bar is 10μm. (F, G)
1256 Representative image of a flushed colon from *Tmu* (F) and *Tc* (G) colonized mice. *Tmu* and *Tc*
1257 stained with anti-*Tmu* antibody (cyan or magenta, respectively), mucus stained with UEA1
1258 (green), nuclei stained with DAPI (blue), epithelium stained with E-Cadherin (white). Scale bars
1259 are 150μm. Each symbol represents an individual mouse (A, B). *p<0.05, **p<0.01, ***p<0.001,
1260 ****p<0.0001, NS, not significant by Students t test.
1261
1262 **Figure S4: Characterization of fermentative metabolic pathways in parabasalids, related**
1263 **to Figure 4.**

1264 (A) Tuft cell frequency in the epithelium of the distal small intestine in WT or *Gpr91*$^{-/-}$ mice
1265 colonized by each protist or fed succinate in the drinking water. (B) Quantification of
1266 extracellular lactate in the cecal contents of mice colonized with each protist for two weeks. (C)
1267 Genome assembly statistics for *Tmu* and *Tc* draft genomes. (D) Short chain fatty acid
1268 quantification in purified protist lysates (E) Heatmap showing the abundance of 48 metabolites
1269 identified in *Tmu* and *Tc* lysates using LC-MS methodologies. Values are plotted as $\log_2$ of the
1270 relative metabolite abundance. (F) Succinate (left) and lactate (right) quantification in culture
1271 supernatants of *P. hominis* (order *Trichomonadida*) and *M. colubrorum* (order
1272 *Tritrichomonadida*). Data is plotted as mean with SD in B and F. Each symbol represents an
1273 individual mouse pooled from two experiments (A). Error bars represent standard deviation (B,
1274 F). *p<0.05, **p<0.01, ***p<0.001, NS, not significant by Student's t test.
1275
1276 **Figure S5. *Tmu* and *Tc* occupy different nutritional niches within the microbiota, related**
1277 **to Figure 5**.
1278 (A) Abundance of *Tmu* and (B) *Tc* in the cecal contents of mice fed diets with different fiber
1279 compositions for two weeks. Fibers in the defined diets are limited to inulin (I), cellulose (C), or
1280 no fiber (None). (C and D) Relative abundance, from 16S rRNA Sequencing, of the 15 most
1281 abundant operational taxonomic units in the distal SI (C) or cecum (D) of mice with or without
1282 each protist. Averages of five mice per group are shown. (E) Relative abundance of *A.*
1283 *muciniphila*, as measured by qPCR, in the stool of mice colonized with or without each protist
1284 for three weeks. (F) Abundance of genes encoding CAZymes in glycoside hydrolase (GH)
1285 families associated with activity on host glycans or plant glycans. White indicates no genes in
1286 the indicated GH family, black indicates greater than or equal to 20 genes. Data is plotted as
1287 mean in (A, B, and E), with each symbol representing an individual mouse. *p<0.05, **p<0.01,
1288 NS, not significant by Mann-Whitney test.
1289

1290 **Figure S6. Fiber deprivation causes *Tmu* to switch to a mucolytic metabolism, related to**
1291 **Figure 6**.
1292 (A) Serial passaging of *Tmu* in PBF medium. Drops in culture density show when *Tmu* was
1293 subcultured into fresh media. (B) Survival of *Tc* in PBF medium after 24 hours, with maltose or
1294 mucus added as the carbon source. (C) *Tmu* abundance in the stool of mice after two weeks of
1295 feeding on def-C chow and different antibiotic treatments. (D) Representative image of *Tmu* in
1296 the colon of a mouse fed def-C chow and fed ampicillin in the drinking water. *Tmu* is stained
1297 with anti-*Tmu* antibody (cyan), mucus is stained with UEA1 (green), nuclei are stained with
1298 DAPI (blue). Scale bar is 100μm. (E) Quantification of *Tmu* localization to the mucus layer in the
1299 colons of mice fed complex chow or def-C chow with ampicillin treatment. Higher mucus/lumen
1300 ratio indicates tighter localization to the colonic mucus layer. (F) Quantification of total bacterial
1301 abundance in the stool of *Tmu*-colonized mice fed def-C chow and given different antibiotic
1302 treatments in the drinking water (mice from Figure 6G). Graphs (A, B, C, and D) depict mean.
1303 Error bars represent standard deviation (B). Each symbol represents an individual mouse (C, E,
1304 F) *p<0.05, **p<0.01, NS, not significant by Mann-Whitney test (C, F) or Students t test (B, E).
1305
1306 **Figure S7. *Tmu* utilizes specific dietary fibers as a carbon source, Related to Figure 7**
1307 (A) *Tmu in vitro* growth curves with different dietary fibers in PBF medium. Dashed lines show
1308 maximum growth over 5 days in PBF medium without a supplemented carbon source. (B) Tuft
1309 cell frequency and (C) Th2 cell frequency in the distal SI of mice colonized with *Tmu* and fed
1310 def-C chow for three weeks. (D) Succinate quantification in the culture supernatants of *Tmu*
1311 cultured with maltose or mucus as a carbon source. Graphs show mean with SD (A, D). Each
1312 symbol represents an individual mouse (B, C). NS, not significant by Student's t test.
1313
1314 **Table S1: Reference parabasalid ITS sequences used for identification of parabasalids in**
1315 **human metagenomic datasets, related to Figure 1.**

Figure 1

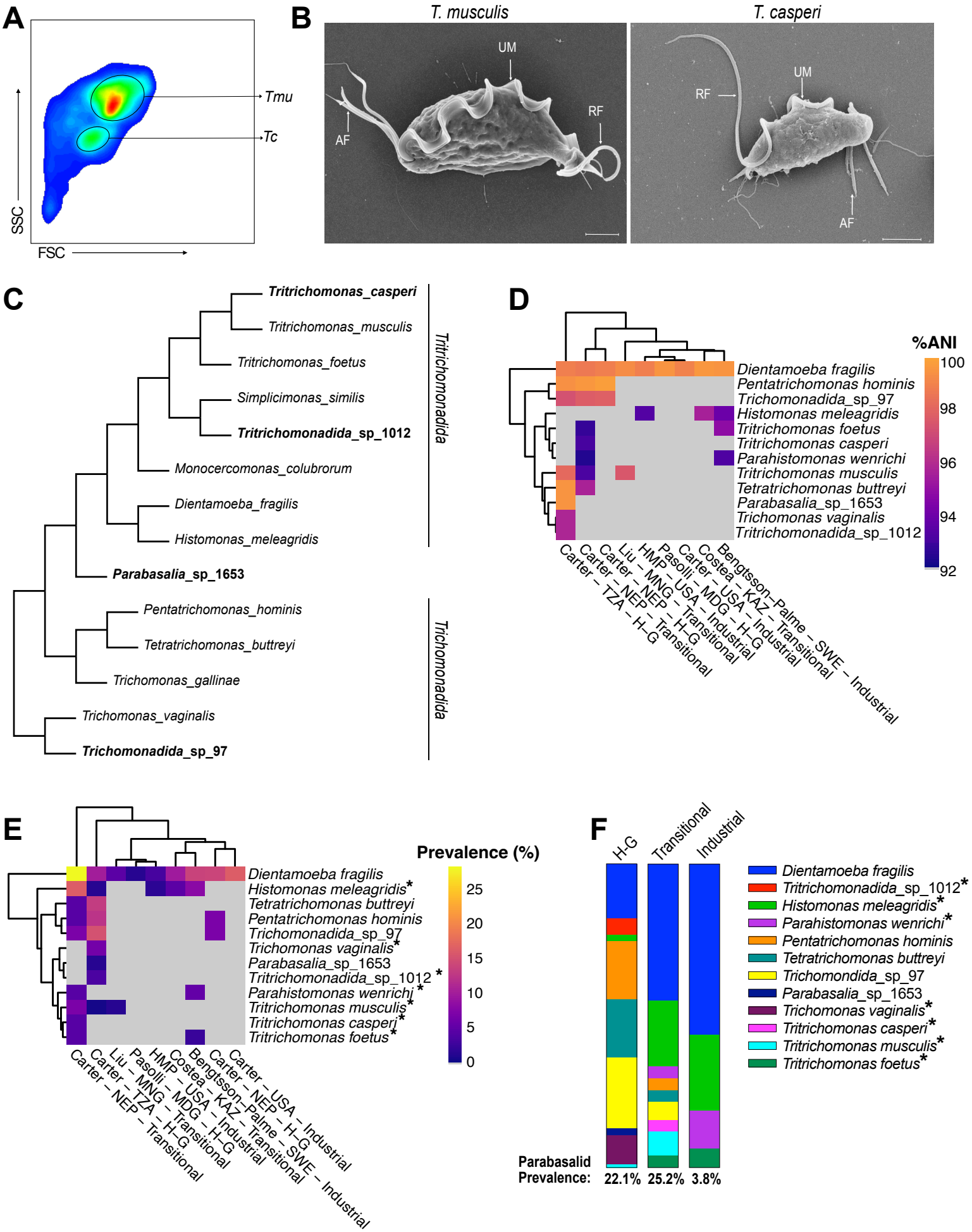

A
Tmu
Tc
SSC
FSC
B
T. musculis
T. casperi
UM
RF
AF
C
Tritrichomonas_casperi
Tritrichomonas_musculis
Tritrichomonas_foetus
Simplicimonas_similis
Tritrichomonadida_sp_1012
Monocercomonas_colubrorum
Dientamoeba_fragilis
Histomonas_meleagridis
Parabasalia_sp_1653
Pentatrichomonas_hominis
Tetratrichomonas_buttreyi
Trichomonas_gallinae
Trichomonas_vaginalis
Trichomonadida_sp_97
Tritrichomonadida
Trichomonadida
D
%ANI
100
98
96
94
92
Dientamoeba fragilis
Pentatrichomonas hominis
Trichomonadida_sp_97
Histomonas meleagridis
Tritrichomonas foetus
Tritrichomonas casperi
Parahistomonas wenrichi
Tritrichomonas musculis
Tetratrichomonas buttreyi
Parabasalia_sp_1653
Trichomonas vaginalis
Tritrichomonadida_sp_1012
Carter – TZA – H–G
Carter – NEP – Transitional
Carter – NEP – H–G
Liu – MNG – Transitional
Carter – USA – Industrial
HMP – USA – H–G
Pasolli – MDG – Industrial
Carter – USA – H–G
Costea – KAZ – Transitional
Bengtsson–Palme – SWE – Industrial
E
Prevalence (%)
25
20
15
10
5
0
Dientamoeba fragilis
Histomonas meleagridis*
Tetratrichomonas buttreyi
Pentatrichomonas hominis
Trichomonadida_sp_97
Trichomonas vaginalis*
Parabasalia_sp_1653
Tritrichomonadida_sp_1012 *
Parahistomonas wenrichi*
Tritrichomonas musculis*
Tritrichomonas casperi*
Tritrichomonas foetus*
Carter – NEP – Transitional
Carter – NEP – H–G
Carter – TZA – H–G
Liu – MNG – Transitional
Carter – USA – H–G
Pasolli – MDG – Industrial
HMP – USA – H–G
Costea – KAZ – Transitional
Bengtsson–Palme – SWE – Industrial
Carter – USA – Industrial
F
H-G
Transitional
Industrial
Parabasalid Prevalence:
22.1%
25.2%
3.8%
Dientamoeba fragilis
Tritrichomonadida_sp_1012*
Histomonas meleagridis*
Parahistomonas wenrichi*
Pentatrichomonas hominis
Tetratrichomonas buttreyi
Trichomondida_sp_97
Parabasalia_sp_1653
Trichomonas vaginalis*
Tritrichomonas casperi*
Tritrichomonas musculis*
Tritrichomonas foetus*

Figure 2

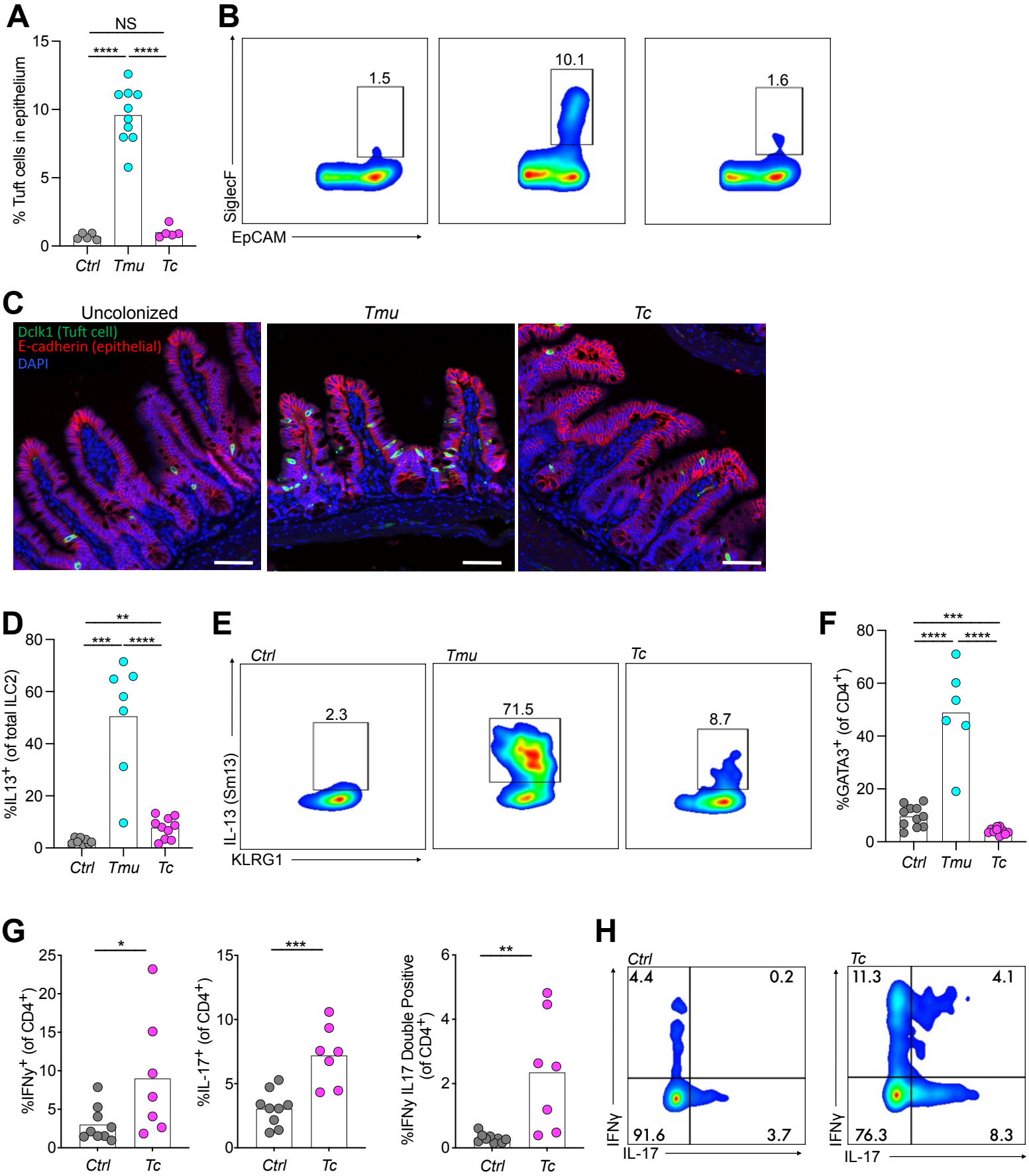

A
NS
****
****
% Tuft cells in epithelium
15
10
5
0
Ctrl
Tmu
Tc
B
1.5
10.1
1.6
SiglecF
EpCAM
C
Uncolonized
Tmu
Tc
Dclk1 (Tuft cell)
E-cadherin (epithelial)
DAPI
D
**
***
****
%IL13+ (of total ILC2)
80
60
40
20
0
Ctrl
Tmu
Tc
E
Ctrl
Tmu
Tc
2.3
71.5
8.7
IL-13 (Sm13)
KLRG1
F
***
****
****
%GATA3+ (of CD4+)
80
60
40
20
0
Ctrl
Tmu
Tc
G
*
%IFNγ+ (of CD4+)
25
20
15
10
5
0
Ctrl
Tc
***
%IL-17+ (of CD4+)
15
10
5
0
Ctrl
Tc
**
%IFNγ IL17 Double Positive (of CD4+)
6
4
2
0
Ctrl
Tc
H
Ctrl
4.4
0.2
91.6
3.7
IFNγ
IL-17
Tc
11.3
4.1
76.3
8.3
IFNγ
IL-17

Figure 3

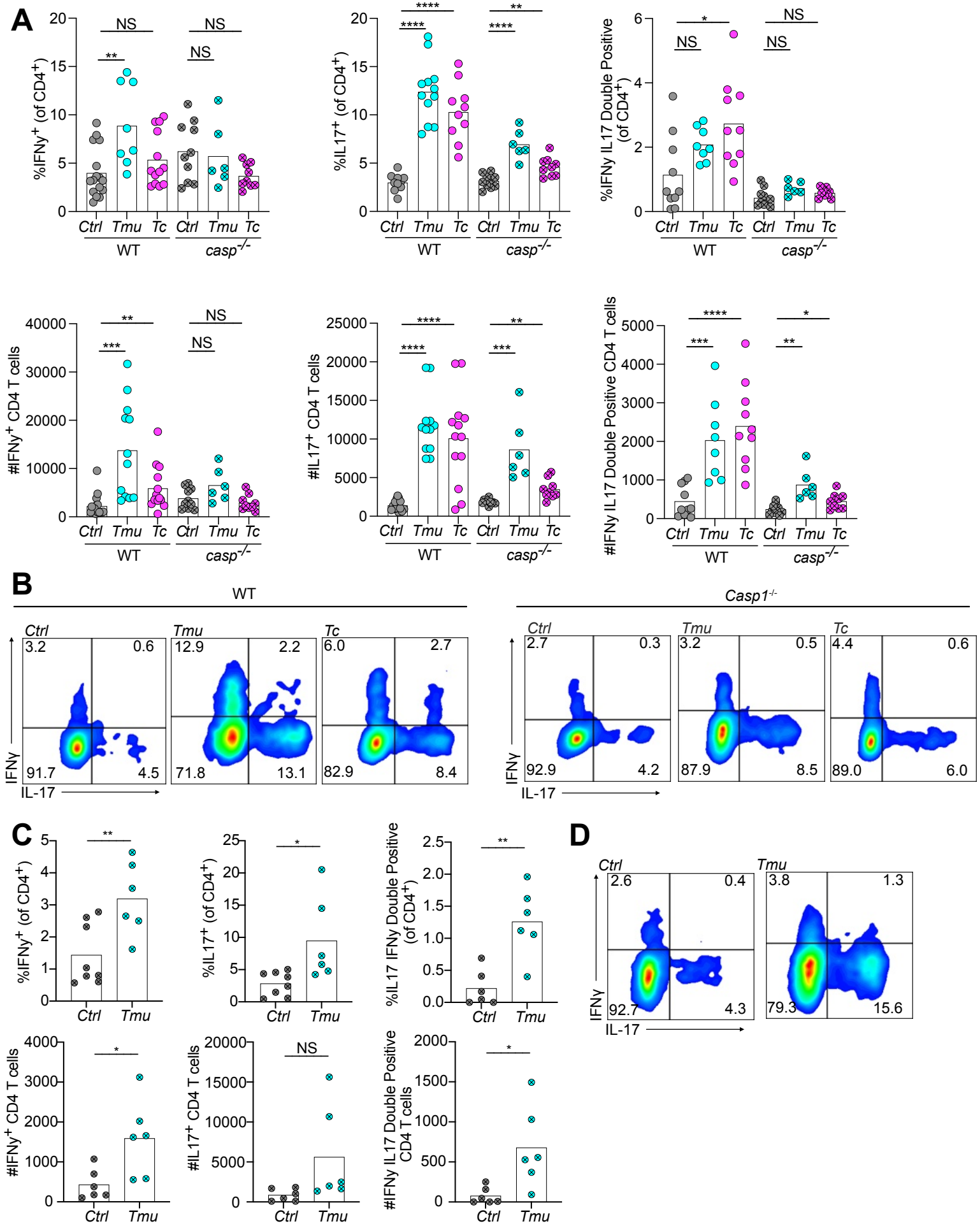

A
%IFNγ+ (of CD4+)
%IL17+ (of CD4+)
%IFNγ IL17 Double Positive (of CD4+)
#IFNγ+ CD4 T cells
#IL17+ CD4 T cells
#IFNγ IL17 Double Positive CD4 T cells
Ctrl Tmu Tc Ctrl Tmu Tc
WT casp-/-
NS
B
WT
Casp1-/-
Ctrl
Tmu
Tc
IFNγ
IL-17
3.2 0.6 91.7 4.5
12.9 2.2 71.8 13.1
6.0 2.7 82.9 8.4
2.7 0.3 92.9 4.2
3.2 0.5 87.9 8.5
4.4 0.6 89.0 6.0
C
D
2.6 0.4 92.7 4.3
3.8 1.3 79.3 15.6

Figure 4

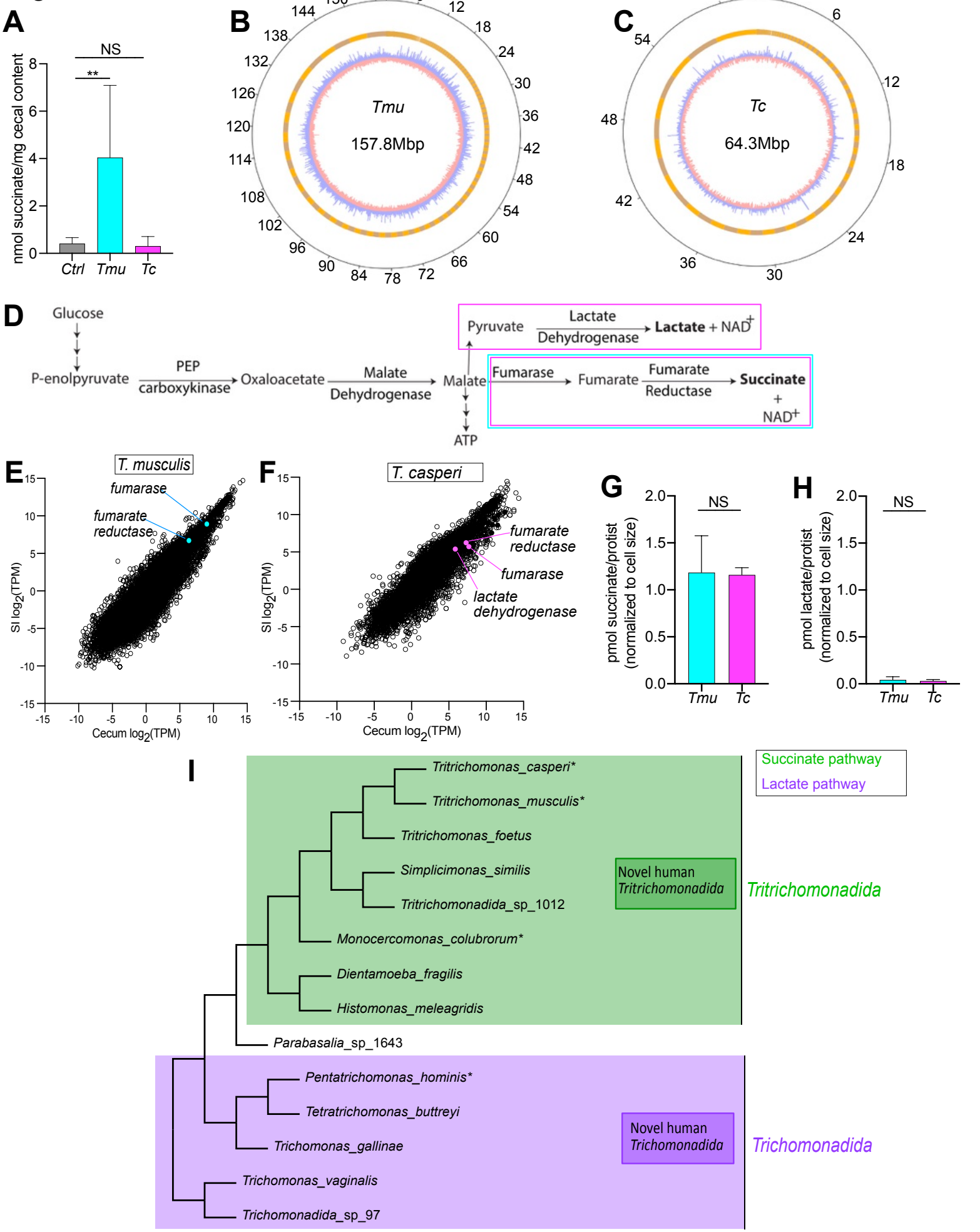

A
NS
**
nmol succinate/mg cecal content
Ctrl
Tmu
Tc
B
Tmu
157.8Mbp
C
Tc
64.3Mbp
D
Glucose
P-enolpyruvate
PEP carboxykinase
Oxaloacetate
Malate Dehydrogenase
Malate
ATP
Pyruvate
Lactate Dehydrogenase
Lactate + NAD+
Fumarase
Fumarate
Fumarate Reductase
Succinate + NAD+
E
T. musculis
fumarase
fumarate reductase
SI log2(TPM)
Cecum log2(TPM)
F
T. casperi
fumarate reductase
fumarase
lactate dehydrogenase
G
pmol succinate/protist (normalized to cell size)
H
pmol lactate/protist (normalized to cell size)
I
Succinate pathway
Lactate pathway
Tritrichomonas_casperi*
Tritrichomonas_musculis*
Tritrichomonas_foetus
Simplicimonas_similis
Tritrichomonadida_sp_1012
Monocercomonas_colubrorum*
Dientamoeba_fragilis
Histomonas_meleagridis
Novel human Tritrichomonadida
Tritrichomonadida
Parabasalia_sp_1643
Pentatrichomonas_hominis*
Tetratrichomonas_buttreyi
Trichomonas_gallinae
Trichomonas_vaginalis
Trichomonadida_sp_97
Novel human Trichomonadida
Trichomonadida

Figure 5

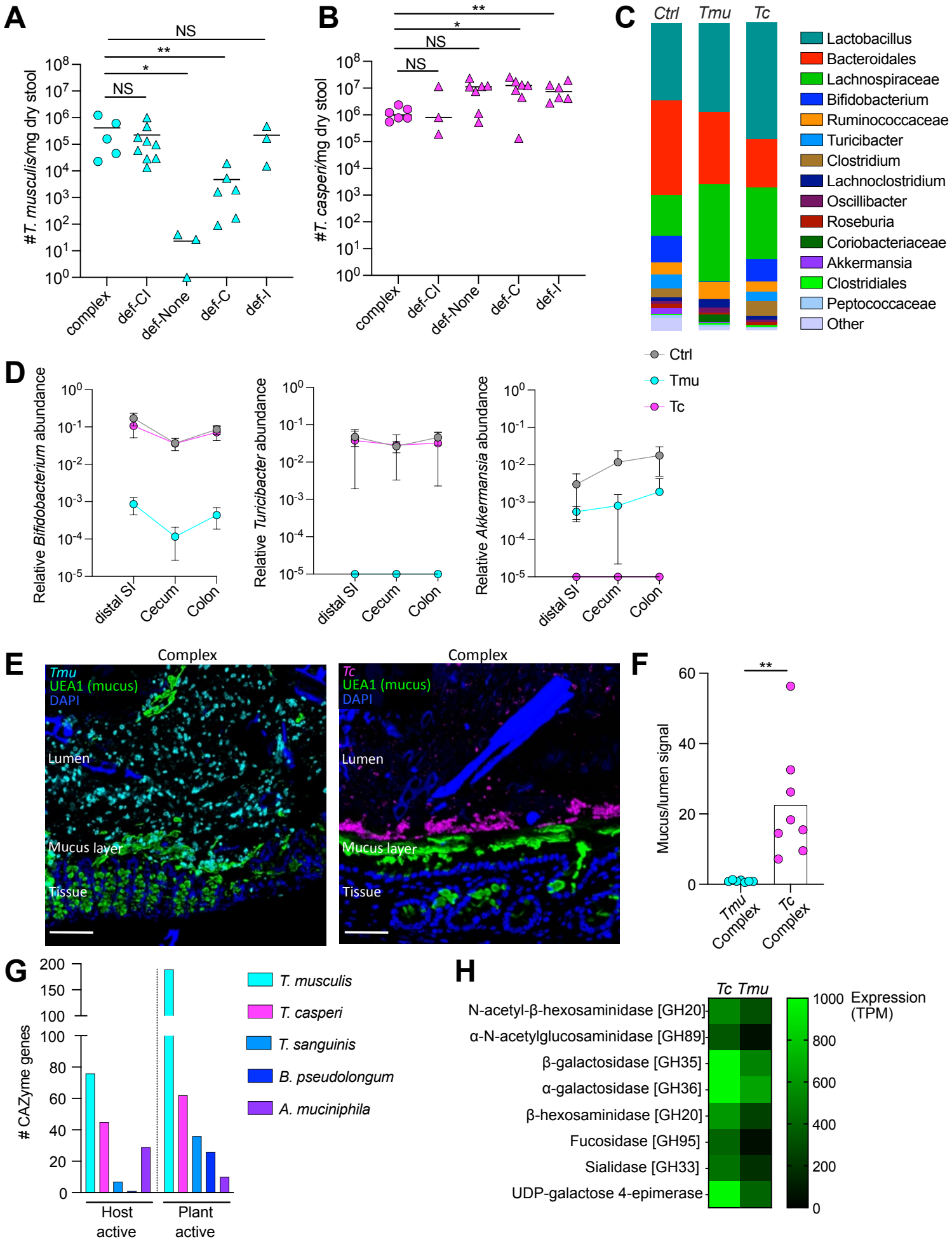

A
NS
**
*
NS
#T. musculis/mg dry stool
complex
def-CI
def-None
def-C
def-I
B
**
*
NS
NS
#T. casperi/mg dry stool
complex
def-CI
def-None
def-C
def-I
C
Ctrl
Tmu
Tc
Lactobacillus
Bacteroidales
Lachnospiraceae
Bifidobacterium
Ruminococcaceae
Turicibacter
Clostridium
Lachnoclostridium
Oscillibacter
Roseburia
Coriobacteriaceae
Akkermansia
Clostridiales
Peptococcaceae
Other
D
Ctrl
Tmu
Tc
Relative Bifidobacterium abundance
Relative Turicibacter abundance
Relative Akkermansia abundance
distal SI
Cecum
Colon
E
Complex
Tmu
UEA1 (mucus)
DAPI
Lumen
Mucus layer
Tissue
Complex
Tc
UEA1 (mucus)
DAPI
Lumen
Mucus layer
Tissue
F
**
Mucus/lumen signal
Tmu Complex
Tc Complex
G
CAZyme genes
T. musculis
T. casperi
T. sanguinis
B. pseudolongum
A. muciniphila
Host active
Plant active
H
Tc Tmu
N-acetyl-β-hexosaminidase [GH20]
α-N-acetylglucosaminidase [GH89]
β-galactosidase [GH35]
α-galactosidase [GH36]
β-hexosaminidase [GH20]
Fucosidase [GH95]
Sialidase [GH33]
UDP-galactose 4-epimerase
Expression (TPM)

Figure 6

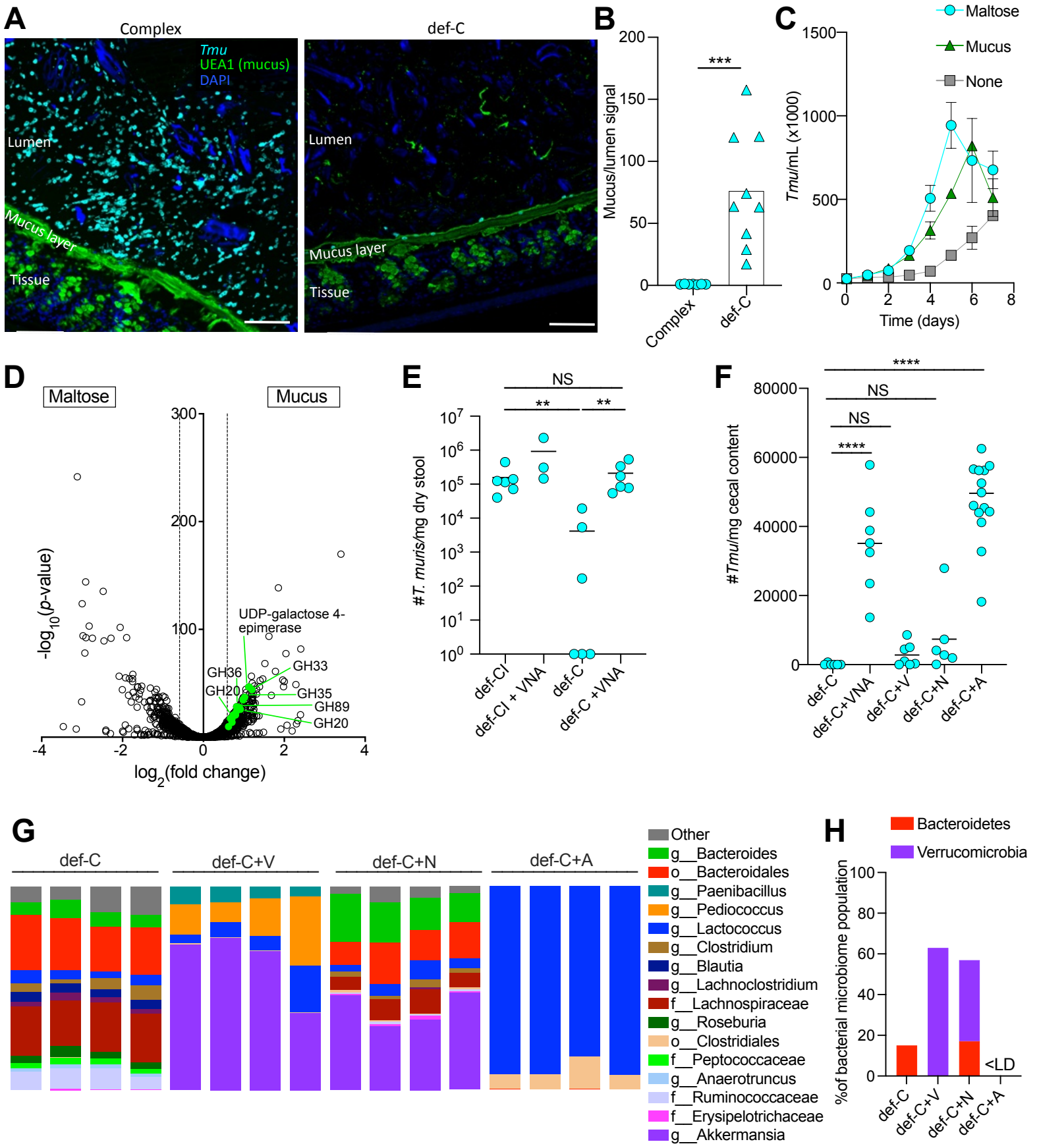
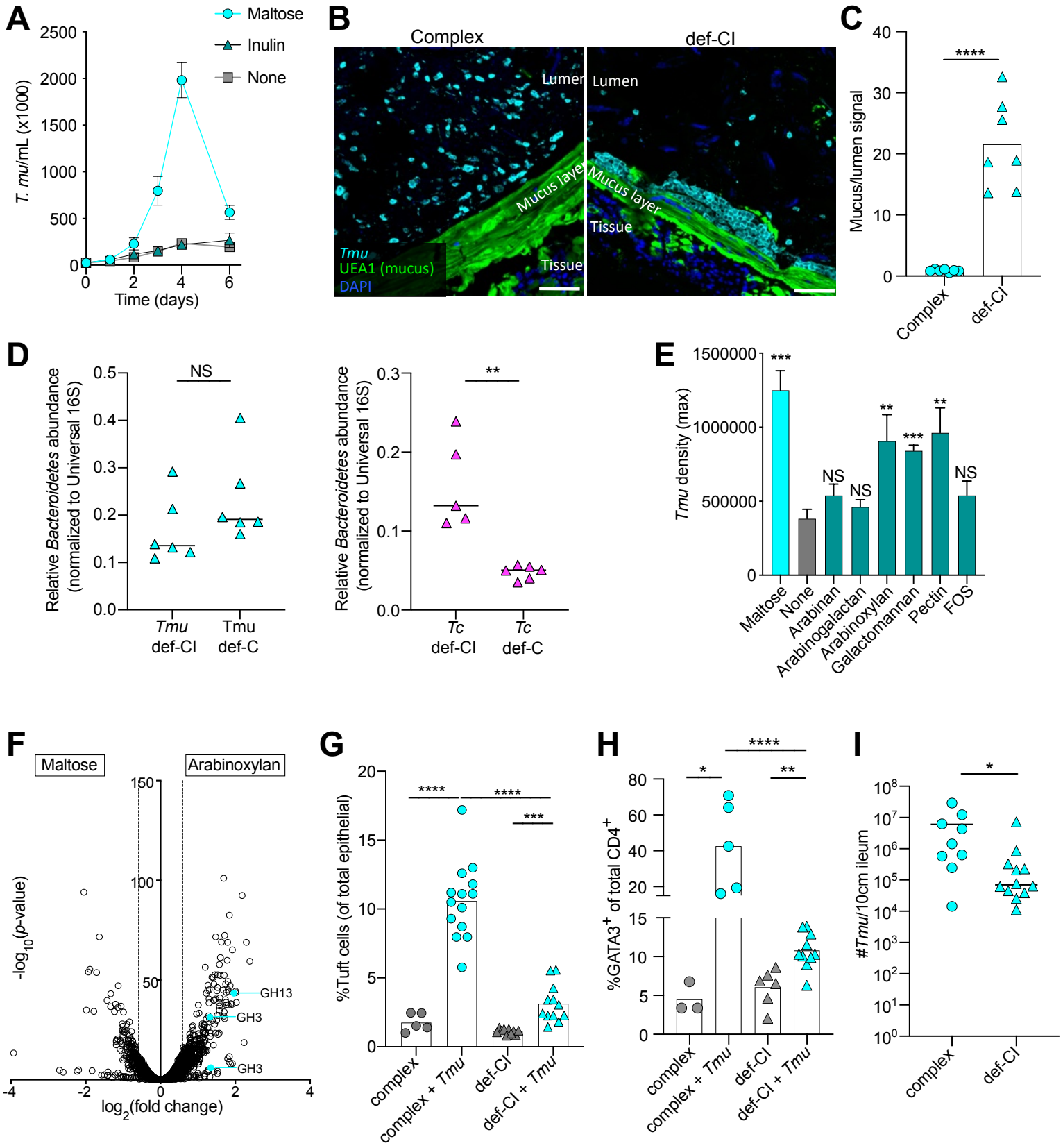

A
Complex
def-C
Tmu
UEA1 (mucus)
DAPI
Lumen
Mucus layer
Tissue
B
Mucus/lumen signal
***
Complex
def-C
C
Maltose
Mucus
None
Tmu/mL (x1000)
Time (days)
D
Maltose
Mucus
-log10(p-value)
log2(fold change)
UDP-galactose 4-epimerase
GH36
GH20
GH33
GH35
GH89
GH20
E
NS
**
**
#T. muris/mg dry stool
def-CI
def-CI + VNA
def-C
def-C +VNA
F
****
NS
NS
****
#Tmu/mg cecal content
def-C
def-C+VNA
def-C+V
def-C+N
def-C+A
G
def-C
def-C+V
def-C+N
def-C+A
Other
g__Bacteroides
o__Bacteroidales
g__Paenibacillus
g__Pediococcus
g__Lactococcus
g__Clostridium
g__Blautia
g__Lachnoclostridium
f__Lachnospiraceae
g__Roseburia
o__Clostridiales
f__Peptococcaceae
g__Anaerotruncus
f__Ruminococcaceae
f__Erysipelotrichaceae
g__Akkermansia
H
Bacteroidetes
Verrucomicrobia
%of bacterial microbiome population
<LD
def-C
def-C+V
def-C+N
def-C+A

Figure 7


A
Maltose
Inulin
None
T. mu/mL (x1000)
Time (days)
B
Complex
def-CI
Lumen
Mucus layer
Tissue
Tmu
UEA1 (mucus)
DAPI
C
Mucus/lumen signal
Complex
def-CI
D
Relative Bacteroidetes abundance (normalized to Universal 16S)
NS
Tmu def-CI
Tmu def-C
Tc def-CI
Tc def-C
E
Tmu density (max)
Maltose
None
Arabinan
Arabinogalactan
Arabinoxylan
Galactomannan
Pectin
FOS
F
Maltose
Arabinoxylan
-log10(p-value)
log2(fold change)
GH13
GH3
GH3
G
%Tuft cells (of total epithelial)
complex
complex + Tmu
def-CI
def-CI + Tmu
H
%GATA3+ of total CD4+
complex
complex + Tmu
def-CI
def-CI + Tmu
I
#Tmu/10cm ileum
complex
def-CI