

1 **Gut Microbiome Composition is Associated with Spatial Structuring and Social interactions in**
2 **Semi-Feral Welsh Mountain Ponies**

3

4 Rachael E. Antwis¹, Jessica M. D. Lea², Bryony Unwin¹, Susanne Shultz²

5 1. School of Environment and Life Sciences, University of Salford, Salford, UK,

6 2. School of Earth and Environmental Sciences, University of Manchester, Manchester, UK,

7

8 **Author Email Addresses:** r.e.antwis@salford.ac.uk; jmd.lea@gmail.com;

9 bryonykate_unwin@hotmail.com; Susanne.shultz@manchester.ac.uk

10

11 **Address for Correspondence:** Dr Rachael Antwis, University of Salford, Room 336, Peel Building,

12 The Crescent, Salford, M5 4WT, UK, 01612954641, r.e.antwis@salford.ac.uk

13

14 **Running Title:** Spatial and social structuring of microbiomes

15 **Keywords:** 16S rRNA gene, amplicon sequencing, harem, horizontal transmission, life-stage, multi-

16 level trait, spatial proximity, social networks, vertical transmission

17

18 **Article Type:** Letter

19 **Number of words in abstract:** 243

20 **Number of words in main text:** 4723

21 **Number of figures:** 4

22 **Number of tables:** 2

23

24

25

26

27 **ABSTRACT**

28 Background: Microbiome composition is linked to host functional traits including metabolism and
29 immune function. Drivers of microbiome composition are increasingly well-characterised; however,
30 evidence of group-level microbiome convergence is limited and may represent a multi-level trait (i.e.
31 across individuals and groups), whereby heritable phenotypes are influenced by social interactions.
32 Here we investigate the influence of spatial structuring and social interactions on the gut microbiome
33 composition of Welsh mountain ponies.

34 Results: Here we show that semi-feral ponies exhibit variation in microbiome composition according to
35 band membership, along with considerable within-individual variation. Spatial structuring was also
36 identified within bands, suggesting that despite communal living, social behaviours still influence
37 microbiome composition. Indeed, we show that specific interactions (i.e. mother-offspring and stallion-
38 mare) lead to more similar microbiomes, further supporting the notion that individuals influence the
39 microbiome composition of one another and ultimately, the group. Foals exhibited different
40 microbiome composition to sub-adults and adults, most likely related to differences in diet.

41 Conclusions: We provide novel evidence that microbiome composition is structured at multiple levels
42 within populations of social mammals and thus, may form a unit on which selection can act. High
43 levels of within-individual variation in microbiome composition, combined with the potential for social
44 interactions to influence microbiome composition, suggest the direction of microbiome selection may
45 be influenced by the individual members present in the group. Although the functional implications of
46 this requires further research, these results lend support to the idea that multi-level selection can act
47 on microbiomes.

48

49

50 **BACKGROUND**

51 All metazoan species harbour complex communities of microorganisms referred to as host
52 microbiomes. The host plus its microbiome complement can be considered as a distinct biological
53 entity, the holobiont, with a complementary genome, the hologenome [1]. Although the concept of a
54 holobiont remains a topic of debate [2,3], there are a several compelling arguments for why this
55 approach is biologically relevant. First, microbiome composition can be heritable such that offspring
56 microbiomes resemble those of their parents [4]. Second, microbiome genomes are much more plastic
57 and evolvable than host genomes over short periods of time, providing a host with phenotypic
58 plasticity that can respond more rapidly to external and internal challenges than the host [5]. Third,
59 these diverse communities are associated with host functional traits, such as immune function and
60 metabolism, as demonstrated across a range of host sites for both mammalian and non-mammalian
61 taxa [6–8]. Indeed, microbiome composition and complexity has been associated with disease
62 prevalence for many host taxa at the individual- and population-level [9–12]. Many host species have
63 a considerable “core” microbiome that is stable with a body site across individuals, time, and space
64 [13–17]. This core microbiome is thought to represent the heritable component of the microbiome that
65 form the “house-keeping” component of the microbiome, while there is also a flexible component to
66 the microbiome that varies based on environmental influences [5]. At the same time, there is
67 significant temporal variation in microbiome composition between and within individuals of a given
68 species [18,19].

69 Intriguingly, group-level microbiome similarity may represent an example of a multi-level trait, where
70 heritable phenotypes are influenced by association patterns [20]. In fact, social structure can lead to
71 adaptive and evolutionary changes within the microbiome and, potentially, the host organism [3]. For
72 example, social pollinators, such as honeybees and bumblebees, share a distinct community of
73 bacteria not identified in solitary bee species [21]. The presence of this distinct microbiome provides
74 social bee species with protection from parasitic infection, and thus confers fitness benefits not
75 present in solitary species lacking this shared microbiome [22]. In communities, interaction patterns
76 defined by social networks can be used to characterise the nature of interactions between individuals.
77 These networks have long been suggested to impact on transmission dynamics of disease in humans
78 [23,24] and animals [25,26], but social interactions should also be associated with the opportunity to
79 share microbiomes and thus, may confer fitness benefits [27].

80 Spatial proximity between individuals has been shown to facilitate microbiome exchange even when
81 social behaviours are minimal. For example, gut microbiomes of solitary North American red squirrels
82 (*Tamiasciurus hudsonicus*; [28]) and mouthpart microbiomes of Phofung river frog tadpoles (*Amietia*
83 *hymenopus*; [29]) are spatially structured. Similarly, gopher tortoise (*Gopherus polyphemus*) gut
84 microbiome composition is determined by geographic proximity of burrows, as well as home ranges
85 and kinship [30]. If spatial proximity promotes microbiome similarity, then social structuring determined
86 by patterns of interaction, association and spatial proximity between individuals provides an ideal
87 mechanism for driving sub-population level patterns in host microbiome communities [31]. The role of
88 social interactions in transmitting pathogens and parasites between individuals is well known,
89 however, such behaviours can also alter and influence the composition of the microbiome [27]. For
90 example, yellow baboon (*Papio cynocephalus*) group membership, social networks and grooming
91 interactions predict taxonomic structure of the gut microbiome even after controlling for the effects of
92 diet, kinship and shared environments [32,33]. Similarly, gut microbiomes of chimpanzees (*Pan*
93 *troglydytes*) are associated with interaction frequency [34] and human milk microbiomes are
94 influenced by size of social network and physical/proximal contact with an infant [35]. Despite growing
95 interest in the role of social interactions in determining gut microbiome composition [34,36], the
96 majority of studies focus on primates.

97 Equids provide an interesting test case for microbiome dynamics at the sub-population level. As
98 hindgut fermenters, the Equidae are particularly reliant on microbial digestion for energy and nutrition
99 [5,37]. Free-ranging horses (*Equus ferus caballus*) form harem bands (i.e. family groups) composed of
100 (usually) one mature stallion, multiple mares and their immature offspring. Found in Snowdonia
101 National Park, the semi-feral Carneddau pony is the closest to a wild unmanaged pony population in
102 the UK [38,39]. They are direct descendants of the wild Welsh mountain pony and are a genetically
103 unique and distinct population, rendering them a high conservation priority [38]. Over 300 individuals
104 exist within smaller scattered bands that form complex social networks [38–40]. Individuals within
105 bands engage in varying levels of affiliative behaviour with conspecifics dependant on various factors
106 such as kinship, age, social status and season. Males are socially central (i.e. well connected), while
107 females are more peripheral and tend to have weak bonds with other mares [40]. Although female
108 relationship strength varies between seasons, their position within the social network is stable across
109 years [40].

110 Here we determine how the gut microbiome of semi-feral ponies from Snowdonia National Park is
111 influenced by spatial structuring, social interactions and kin relationships. Using social network
112 analysis combined with 16S rRNA gene amplicon sequencing of faecal samples, we test the following
113 hypotheses; i) there will be within-individual variation in microbiome composition, but this will not be as
114 large as between-individual variation; ii) mares will have more similar microbiomes to band stallions
115 than to other mares in their band; iii) mares will have more similar microbiomes to their own offspring
116 than to other juveniles in the band; iv) band, life-stage and sex will influence microbiome composition;
117 v) band-level variation in microbiome composition will be driven by spatial structuring (i.e. social
118 networks).

119

120 **METHODS**

121 *Study Animals*

122 Carneddau Welsh mountain ponies are located in the Carneddau mountain range, Snowdonia
123 National Park, North Wales (53.22°N, 3.95°W) over an area of approximately 35-40 km² of commons
124 land between 287 and 610 m above sea level. The land is used primarily for sheep farming and
125 recreational hiking and thus, ponies are habituated to human presence but not to physical contact.
126 The population is essentially unmanaged aside from an annual roundup event in November, during
127 which individuals are herded onto adjacent farmland for one to two days for population management
128 purposes. Individuals can be identified using their age-sex classification and a photographic database
129 that depicts coat colour, face and leg markings and ear tags/notches. For this study, we collected data
130 from 30 individuals across three focal bands (Aber, Marsh and Valley) that have been the subjects of
131 long-term behavioural and demographic data collection [39,40] (Table 1).

132

133 *Distribution Mapping and Social Network Analysis*

134 Demography and proximity data were collected over 10 sampling days between the 21st August and
135 14th November 2014 (the same time period when faecal sampling also occurred). All ponies included
136 in the spatial analyses were sighted a minimum of five days, sampled opportunistically within the study
137 area. Upon encountering a group, we recorded time, pony IDs and GPS location along with an

138 approximate spatial network of the ponies. We plotted the geographic distribution of the bands over
139 the study period using the ggmap package [41] in RStudio (v1.0.153) [42] for R (v3.4.1) [43].
140 We approximated the distance in metres between individuals. All individuals less than ~100 m apart
141 and moving as a cohesive unit were considered to be associated with each other [40]. Association
142 matrices were constructed for each day of sampling; individuals that were close together (< 15 m) or
143 interacted were given a score of 2, other individuals (i.e. those 15 – 100 m apart) were given a score
144 of 1 and more than 100 m apart scored 0. Using these association scores, an overall weighted
145 association index for each dyad was calculated using a modified version of the simple ratio index [44],
146 where edge weight was calculated as:

147
$$E_{AB} = \frac{x_{SUM}}{2x_{COUNT} + y_{AB} + y_A + y_B}$$

148
149 where x_{SUM} is the sum of associations between individuals A and B , x_{COUNT} is the number of times A
150 and B have been sighted together (where x_{COUNT} multiplied by two is the maximum possible
151 association score), y_{AB} is the number of times both A and B were observed but not together, y_A is the
152 number of times only individual A was seen and y_B is the number of times only B was seen.

153

154 *Sample Collection and 16S rRNA Gene Amplicon Sequencing*

155 For each band, faecal samples were collected from the stallion plus 4-7 mares and 2-5 juveniles
156 (Table 1 and Table S1) between the 21st August and 11th November 2014, prior to the annual round
157 up. Faecal samples were collected using sterile gloves. Most samples were collected within 10
158 minutes of defecation, but on rare occasions, this took up to a maximum of one hour when multiple
159 individuals defecated within a short period. Several samples were collected from different parts of the
160 dung pile, but no faeces in contact with the ground was collected. Thus, there was minimal risk of
161 environmental contamination. The samples were mixed thoroughly by hand in a sterile bag and a
162 subsample retained for analysis. Three to five samples were collected per individual across the four
163 study months (Table S1). Samples were stored and transported in cool bags to the University of
164 Manchester the same day and frozen at -80°C prior to DNA extraction.

165 DNA was extracted using the QIAamp DNA Stool Mini Kit (Qiagen, UK) following the manufacturer's
166 protocol with an additional incubation time of 30 minutes at 95°C. A blank extraction was also included

167 to act as a negative control for sequencing. DNA was amplified for the 16S rRNA gene (v4 region)
168 using dual indexed forward and reverse primers according to Kozich et al. [45] and Griffiths et al. [29].
169 Briefly, PCRs were run in duplicate using Solis BioDyne 5x HOT FIREPol® Blend Master Mix, 2µM
170 primers and 1µl of sample DNA. Thermocycling conditions were as follows: 95°C for 15 minutes; 28
171 cycles of (95°C for 20s, 50°C for 60s, 72°C for 60s), and a final extension at 72°C for 10 minutes. PCR
172 replicates were checked on an Agilent 2200 TapeStation, combined into a single PCR plate and
173 cleaned using HighPrep™ PCR clean up beads (MagBio, USA) according to the manufacturers'
174 instructions. Products were quality checked using an Agilent 2200 TapeStation and quantified using a
175 Qubit™ 3.0 Fluorometer according to the manufacturers' protocol. Samples were pooled according to
176 concentrations in order to minimise sequencing bias. Paired-end (2 x 250bp) amplicon sequencing
177 was conducted on an Illumina MiSeq platform with negative and positive (mock community) controls.

178

179 *Pre-Processing of Microbiome Data*

180 We conducted all analyses in RStudio (v1.0.153) [42] for R (v3.4.1) [43]. A total of 3,208,334 raw
181 sequence reads from 112 samples were generated during sequencing. We conducted sequence
182 processing in dada2 v1.5.0 [46] using the default pipeline (see Supplementary Information). Modal
183 contig length was 253bp once paired end reads were merged. We removed sequence variants (SVs)
184 with length >260bp (4 SVs; 0.086% of total sequences) along with chimeras and two SVs found in the
185 negative controls, leaving an average of 22294 reads per sample (range 8071 – 42869). We assigned
186 taxonomy using the SILVA v128 database [47,48]. To provide greater taxonomic detail about
187 unidentified SVs and to stop the removal of these during analyses that agglomerate to a given
188 taxonomic level, we fully annotated the taxonomy table to species level using higher levels
189 assignments (e.g. SV1 was named "Family_Prevotellaceae" at the genus and species levels). We
190 exported the final SV table, taxonomy table and sample metadata to the phyloseq package [49] and
191 converted the data to relative abundance for further analyses.

192

193 *Microbiome Variation According to ID*

194 We produced an NMDS plot in phyloseq using the Bray-Curtis distance matrix to visualise the
195 variation within and between individuals according to community composition. To determine the

196 microbiome variation attributable to individual variation (ID), we conducted a permutational ANOVA
197 (PERMANOVA; adonis) in the vegan package [50].

198 We calculated the core microbiome of individual samples using a detection threshold of 0.001% and a
199 prevalence threshold of 99.9% (i.e. a given SV must be present in 99.9% of individuals with a relative
200 abundance of at least 0.001%) in the microbiome package [51]. We used an NMDS plot to visualise
201 the variation in core microbiome according to ID and analysed the data using an adonis analysis (as
202 above).

203 To determine whether there was greater microbiome variation *within* an individual than *between*
204 individuals, we calculated Jensen-Shannon Divergence (JSD) values in the phyloseq package [49].
205 JSD values give a measure of similarity between all individual samples (i.e. by calculating the distance
206 between samples) either from the same individual (i.e. within-individual variation) or from different
207 individuals in the same band (i.e. between-individual variation). Smaller JSD values indicate more
208 similar microbial communities and conversely, larger values indicate a less similar community. We
209 used a generalised linear mixed model with ID and band as random factors to compare JSD distances
210 within individuals to JSD distances between individuals and visualised the data using a box plot.

211

212 *Microbiome Variation According to Band, Life-Stage and Sex*

213 We categorised individuals under the age of 1 year as foals; for older individuals, females <2 years old
214 and males <3 years old as sub-adults (females are usually reproductively mature from 2 years
215 onwards but males take longer to mature, disperse and attract mares). We classified all others as
216 adults, with the exception of one female who still displayed sub-adult behaviours, did not disperse
217 from her natal band and had not foaled by age 3, and so was considered a sub-adult. We visualised
218 the taxonomic composition (at the class level) of the communities according to band and life-stage
219 using stacked plots in phyloseq [49] and ggplot2 [52].

220 To obtain the “average microbiome” for an individual, we merged raw sample data within an individual
221 using the merge_samples function in phyloseq (using “fun=mean”) [49]. To determine whether there
222 was greater microbiome variation between bands than within bands, we calculated Jensen-Shannon
223 Divergence (JSD) values between individuals, using data from their average microbiome, as described

224 above. We used a one-way ANOVA with Tukey's posthoc analysis to compare JSD distances within
225 and between bands, and visualised the data using a box plot.

226 We produced NMDS plots in phyloseq using the Bray-Curtis distance matrix to visualise differences in
227 beta diversity according to band and life-stage. We used an adonis analysis to test for significant
228 effects of band, life-stage and sex on total microbiome community composition. We then calculated
229 the core microbiome as described above and repeated the adonis analysis for this core community.
230 Additionally, we agglomerated the core taxa to genus level and visualised the core microbiome as a
231 heat map to give a representation of the bacterial taxa present.

232 To identify differences in microbiome composition between foals (which at approximately 5-8 months
233 old, were most likely still nursing) and sub-adults (which were most likely fully weaned), we conducted
234 an indicator analysis using the multipatt function in the indicpecies package [53].

235

236 *Effects of Spatial Structuring and Social Interactions on the Microbiome*

237 We correlated the social network association matrix with the NMDS scores of each individuals'
238 average microbiome using a Kendall's correlation coefficient for non-parametric data with ties.

239 Networks were constructed and visualised using the igraph package [54] with edges weighted by
240 either microbiome similarity (the inverse of the NMDS distance) or the association index as described
241 above. As the microbiome distance matrix is fully connected, we deleted edges with a similarity less
242 than the mean value for the population. We calculated JSD values between merged samples in the
243 phyloseq package [49] and used general linear mixed models (with ID and band as random factors) to
244 identify whether mares had more similar microbiomes to other mares within the same band or to the
245 band stallion; and whether mares had more similar microbiomes to their own offspring than to other
246 mares' offspring within the same band.

247

248 **RESULTS**

249 Bacteria primarily belonged to the Bacteroidia, Clostridia, Spirochaetes and Fibrobacteria classes
250 (Bacteroidetes, Fibrobacteres, Firmicutes and Spirochaetae phyla) (Figs. S1 and S2). The dominant
251 families represented in the core microbiome were anaerobic bacteria associated with grass-eating

252 mammals, including Prevotellaceae, Ruminococcaceae, Rikenellaceae, Lachnospiraceae,
253 Spirochaetaceae, Fibrobacteraceae, Christensenallaceae, Erysipelotrichaceae, Acidaminococcaceae
254 and various groups of Bacteroidales.

255 An adonis analysis showed pony ID had a significant effect on total microbiome composition ($p <$
256 0.001 ; Table 2), with 52.6% of the variation in the microbiome attributable to individual variation (Fig.
257 1). We obtained similar results for the adonis with the core microbiome ($p < 0.001$; Table 2), with
258 49.6% of the variation explained by ID. Despite the large amount of microbiome variation explained by
259 ID, within-individual samples had significantly lower JSD values (mean of 0.255 ± 0.006) than
260 between-individual samples (mean of 0.347 ± 0.001) ($X^2 = 391.62$, d.f. = 1, $p < 0.001$) (Fig. 2a). That
261 is, there is greater variation between individuals (average of 35%) than within individuals (average of
262 26%).

263 There were significant effects of band and life-stage on total microbiome composition, but not sex
264 (Table 2, Figs. 3a and 3b). In both cases, the proportion of the variation in the microbiome for these
265 significant factors (14.0% for band and 10.4% for life-stage) was much lower than for pony ID, as
266 analysed previously (52.6%). The results of the adonis analysis for the core microbiome were similar
267 to those for the total microbiome, where band and life-stage both significantly affected core
268 microbiome composition, but not sex (Table 2). Band and life-stage account for a slightly larger
269 proportion of the variation in the core microbiome (19.4% and 16.6%, respectively) than the total
270 microbiome.

271 Consistent with the spatial distribution of the bands (Fig. 4a), microbiome composition of individuals in
272 Valley differed considerably to those in Aber and Marsh, which are more similar to each other but still
273 display some degree of separation (Fig. 3a). There was a significant difference in JSD metric values
274 within and between the bands ($F_{5,424} = 6.557$, $p < 0.001$), and the Tukey posthoc indicated that within-
275 band variation for Valley was significantly lower than the variation within the other two bands, and
276 significantly lower than between-band variation for all three combinations (Figure 2b). In addition to
277 this band-level differentiation of microbiomes, there was a significant correlation between social
278 network tie weight (i.e. spatial distribution) and microbiome composition ($T = -0.11$, $p < 0.001$) within
279 bands, such that individuals that associate more have more similar microbiomes (Fig 4b and 4c).

280 The microbiome of foals was considerably different to that of sub-adults and adults, whereas these
281 latter two groups were very similar to one-another (Fig 3b and S2). An indicator analysis identified six
282 bacterial genera (out of a possible 188) that were significantly associated ($p < 0.05$) with sub-adults
283 compared with foals; *Prevotellaceae Ga6A1 group*, *Denitrobacterium*, *Oscillibacter*, *Anaerovibrio*,
284 *Family_CR-115*, and *Anaerostipes*. There were no genera significantly associated with foals
285 compared with sub-adults; that is, there were no genera uniquely associated with foals compared to
286 sub-adults.

287 Maternal relationship had a significant effect on microbiome similarity ($X^2 = 8.425$, d.f. = 2, $p = 0.015$;
288 Fig. 2c). Pairwise comparisons showed significant differences in microbiome divergence between
289 mother-offspring relationships and non-maternal mares and juveniles (foals and sub-adults combined)
290 ($p = 0.017$). Microbiome divergences between foal-less mares and juveniles were not significantly
291 different to those of mother-offspring relationships ($p = 0.313$) or non-maternal mares and juveniles (p
292 = 1.000). Mares had significantly more similar microbiomes to the band stallion than the other mares
293 in their band ($X^2 = 4.206$, d.f. = 1, $p = 0.040$; Fig. 2d).

294

295 **DISCUSSION**

296 The effects of population structuring in general, and social interactions in particular, on microbiome
297 composition remains poorly understood; challenges often arise in the separation of direct microbiota
298 transmission via social interactions from effects of communal living such as a shared diet or physical
299 environment [27,32]. Here, we show that despite large variation between individuals in microbiome
300 composition, spatial structuring, social relationships (i.e. mother-offspring and stallion-mare) and
301 network ties account for microbiome similarities. The main predictor of microbiome composition is
302 individual identity (pony ID) accounting for around 50% of microbiome variation, with up to 26%
303 variation across multiple samples collected for each individual and up to 34% variation between
304 individuals. Significant inter-individual variation in microbiome composition has been shown in other
305 species [18,19,28,55–57]. Given that such a large component of microbiome variation is due to
306 individual ID, which individuals are present within the band (both mares and stallions) will likely
307 influence the composition of the total group microbiome.

308 Band membership also predicted microbiome composition, with ~14% of the total microbiome
309 variation and ~19% of the core microbiome variation explained by this factor. That the microbiome
310 composition of ponies belonging to Aber and Marsh are more similar to one another than Valley may
311 be driven by both spatial structuring and diet, given that the home ranges of these bands overlap. The
312 home ranges of Aber and Marsh are also somewhat different to that of Valley in terms of elevation,
313 slope and soil moisture; these are more low-lying and marshier in comparison to the steeper, more
314 well drained and exposed slopes that characterised the home range of Valley during the study period.
315 The type and quality of grasses or forage across the study area (approximately 5km²) also vary
316 according to habitat type and thus, diet quality may be driving the observed differences in bands. In
317 addition, variation in browsing behaviour may be driving differences in microbiome composition
318 between individuals. Dietary composition has been shown to affect the microbiome of vertebrates with
319 consequences for microbiome function and fitness traits such as reproductive success [37,56,58–60].
320 The microbiome of Equidae is highly susceptible to changes in diet with consequences for nutrient
321 assimilation [37] and diet can have a significant effect on population performance [61]. Microbes
322 acquired from the environment (horizontal transfer) are likely to have greater genomic variation than
323 vertically-transmitted symbionts and thus may provide greater variation for microbiome-derived
324 functional advantages [5]. Thus, spatial variation in microbial communities between sub-populations
325 may have implications for fitness traits [5]. Aber and Marsh also showed higher within-band variation,
326 comparable in magnitude to between-band variation, whereas Valley had significantly lower
327 microbiome variation within the band. This may reflect the spatial and environmental differences
328 experienced by members of Valley compared with Aber and Marsh, as well as fewer interactions
329 between Valley and the other two bands. This lower microbiome variation across the group as a whole
330 may have implications for group-level fitness. More work is required to understand how group-level
331 microbiome variation relates to population resilience [27].

332 Although it may be difficult to dissociate between the influence of shared living and diet on microbiome
333 composition *between* bands, spatial structuring was also identified *within* bands, suggesting that
334 despite communal living, social behaviours still influence microbiome composition. Social behaviours,
335 such as grooming, that occur between members of the same band provide an opportunity for
336 individuals to share microbial communities. Moreover, close spatial proximity also promotes the
337 sharing of gut microbiomes through contact with recently deposited faeces, including potential

338 coprophagy [62]. Thus, microbiomes of individuals with close social ties are more likely to converge
339 and indeed, our data show that specific interactions (i.e. mother-offspring and stallion-mare) lead to
340 more similar microbiomes. This further supports the notion that individuals influence the microbiome
341 composition of one another and ultimately, the group. Affiliative behaviours occur more frequently
342 between mothers and their offspring than between foals and non-maternal mares, but vertical
343 transmission of microbiomes between mothers and their foals may also derive from birth as well as
344 transfer of milk during nursing [35,63,64]. Ren et al. [28] also found that microbiomes of mothers and
345 offspring were more similar to one another than between unrelated individuals in red squirrels.
346 Interestingly, foal-less mares had an intermediate microbiome similarity to foals compared with
347 mothers and non-maternal mares, suggesting greater levels of affiliative behaviour or social interaction
348 between foals and mares that did not have offspring in the band. Stallions occupy a central social role
349 in the group, unlike less well-connected mares [40], which is reflected in the greater microbiome
350 similarity between stallions and mares (than between mares) as demonstrated here. However, it is not
351 clear whether the convergence of microbiomes is driven by the stallion or the mare, but it may well be
352 both. This may result from affiliative behaviours between stallions and mares (including mating) but
353 may also reflect the behaviour of stallions to smell, and thus come into contact with, mares' faeces.
354 Given that juveniles are prone to dispersal [39,40] and that social structures tend to break up and
355 reform after significant events such as the annual round-up (Lea & Shultz, unpublished data), it would
356 be interesting to follow changes in individuals' microbiomes over such events to determine how quickly
357 these converge and whether microbial signatures of the original band remain. It would also be of
358 interest to compare the microbiome composition of males in bachelor groups to those of stallions to
359 further determine the propensity for mares to alter stallion microbiomes.

360 Although we can estimate similarity between the microbial communities across bands and individuals,
361 we do not yet know how this relates to functional variation or fitness proxies at the sub-population (i.e.
362 band) level. Genetic determinants of microbiome composition and thus, heritability of microbiomes,
363 have been demonstrated across a range of host taxa [4,29,30,57]. That band-level differences in
364 microbiome composition were also significant in the core microbiome further supports the notion that
365 group-level selection may occur within host microbiomes. However, dispersal of individuals between
366 bands means that sub-populations are not genetically-isolated. To further understand the potential for
367 microbiome to act as a unit that selection can act on, it would be valuable to quantify the relative

368 contributions of genetic, environmental and social factors that determine microbiome composition
369 within this system (and across a range of hosts) and to link these to fitness outcomes such as
370 reproductive success and disease susceptibility.

371 We also demonstrate differences between life-stages in microbiome composition of Carneddau
372 ponies; foals had considerably different microbiome composition to both sub-adults and adults. Similar
373 changes in microbiome composition across host development have been seen in other host
374 organisms [10,29,65]. For mammals, this is particularly evident for nursing young compared with
375 weaned individuals [66–68], and this most likely explains the results we see in our data. There was an
376 absence of specific genera in the microbiome of foals, indicating the transition to a grass-based diet
377 leads to the assimilation of additional bacterial groups into the gut microbiome, potentially through
378 environmental transmission. Although gut microbiome composition has been shown to differ between
379 sexes [69], we found that microbiomes were not significantly different between males and females for
380 this population of semi-feral ponies. However, this may reflect a low number of males in the analysis. It
381 would be of interest to follow changes in male microbiome across dispersion, and particularly shifts in
382 composition as stallions' form new family groups and their microbiome is influenced by, and
383 influences, new mares joining their band.

384

385 **CONCLUSIONS**

386 Here we show that semi-feral ponies exhibit variation in microbiome composition between bands,
387 which may relate to social, dietary and environmental factors. In addition, due to the high level of
388 within-individual variation, the direction of selection may be influenced by the individual members
389 present in the group. Spatial structuring was also identified within bands, suggesting that despite
390 communal living, social behaviours still influence microbiome composition. We identify two such
391 interactions; mother-offspring and stallion-mare, that lead to more similar microbiomes, indicating that
392 individuals influence the microbiome composition of one another and ultimately, the group. Thus, we
393 provide novel evidence that microbiome composition is structured at multiple levels within populations.
394 The functional implications of this requires further research.

395

396 **DECLARATIONS**

397 **Ethics Approval**

398 This study was approved by the University of Salford Research, Innovation and Academic
399 Engagement Ethical Approval Panel (ST1617-83) and the University of Manchester (Cat-D; non-
400 licensed procedure).

401

402 **Consent for Publication**

403 Not applicable.

404

405 **Availability of Data and Materials**

406 The datasets generated and analysed for this study are available in the NCBI SRA repository under
407 BioProject PRJNA478495 (<https://www.ncbi.nlm.nih.gov/bioproject/478495>), SRA accession number
408 SRP151639. All R code is available as Rmd files in Supplementary Information.

409

410 **Competing Interests**

411 The authors declare that they have no competing interests.

412

413 **Author Contributions**

414 RA, JL and SS conceived the study; JL and RA collected the samples; BU and RA conducted the lab
415 work; RA, JL and SS analysed the data; BU, SS, JL and RA wrote and revised the paper. All authors
416 read and approved the final manuscript.

417

418 **Funding**

419 SS is supported by a Royal Society University Research Fellowship. JMDL was supported by a
420 University of Manchester faculty studentship.

421

422 **Acknowledgements**

423 The authors would like to thank BEI Resources (Human Microbiome project) for supplying the mock
424 community used in this study.

425

426 REFERENCES

427 1. Bordenstein SR, Theis KR. Host biology in light of the microbiome: Ten principles of holobionts and
428 hologenomes. *PLoS Biol.* 2015;13:1–23.

429 2. Douglas, Angela E W. Holes in the Hologenome : Why Host-Microbe Symbioses Are Not
430 Holobionts. *MBio.* 2016;7:1–7.

431 3. Rosenberg E, Zilber-Rosenberg I. The hologenome concept of evolution after 10 years.
432 *Microbiome.* 2018;6:78.

433 4. Blekhman R, Goodrich JK, Huang K, Sun Q, Bukowski R, Bell JT, et al. Host genetic variation
434 impacts microbiome composition across human body sites. *Genome Biol.* 2015;16:1–12.

435 5. Shapira M. Gut Microbiotas and Host Evolution : Scaling Up Symbiosis. *Trends Ecol Evol.*
436 2016;31:539–549.

437 6. Fitzpatrick CR, Copeland J, Wang PW, Guttman DS, Kotanen PM, Johnson MTJ. Assembly and
438 ecological function of the root microbiome across angiosperm plant species. *Proc Natl Acad Sci.*
439 2018;201717617.

440 7. Hanning I, Diaz-Sanchez S. The functionality of the gastrointestinal microbiome in non-human
441 animals. *Microbiome.* 2015;3:51.

442 8. Morgan XC, Kabakchiev B, Waldron L, Tyler AD, Tickle TL, Milgrom R, et al. Associations between
443 host gene expression, the mucosal microbiome, and clinical outcome in the pelvic pouch of patients
444 with inflammatory bowel disease. *Genome Biol.* 2015;16:1–15.

445 9. McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, Liu Z, et al. The macaque gut microbiome
446 in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathog.* 2008;4.

447 10. Arrieta M-C, Stiemsma LT, Amenyogbe N, Brown EM, Finlay B. The Intestinal Microbiome in Early
448 Life: Health and Disease. *Front Immunol.* 2014;5:1–18.

449 11. Bates KA, Clare FC, O'Hanlon S, Bosch J, Brookes L, Hopkins K, et al. Amphibian

450 chytridiomycosis outbreak dynamics are linked with host skin bacterial community structure. *Nat*
451 *Commun.* 2018;9:1–11.

452 12. Jack ALH, Nelson EB. A seed-recruited microbiome protects developing seedlings from disease
453 by altering homing responses of *Pythium aphanidermatum* zoospores. *Plant Soil.* 2018;422:209–22.

454 13. Turnbaugh PJ, Hamady M, Yatsunencko T, Cantarel BL, Duncan A, Ley RE, et al. A core gut
455 microbiome in obese and lean twins. *Nature.* 2009;457:480–5.

456 14. Huse SM, Ye Y, Zhou Y, Fodor AA. A core human microbiome as viewed through 16S rRNA
457 sequence clusters. *PLoS One.* 2012;7:1–12.

458 15. Kueneman JG, Parfrey LW, Woodhams DC, Archer HM, Knight R, McKenzie VJ. The amphibian
459 skin-associated microbiome across species, space and life history stages. *Mol Ecol.* 2014;23:1238–
460 50.

461 16. Loudon AH, Woodhams DC, Parfrey LW, Archer H, Knight R, McKenzie V, et al. Microbial
462 community dynamics and effect of environmental microbial reservoirs on red-backed salamanders
463 (*plethodon cinereus*). *ISME J.* 2014;8:830–40.

464 17. Laforest-Lapointe I, Messier C, Kembel SW. Host species identity, site and time drive temperate
465 tree phyllosphere bacterial community structure. *Microbiome.* 2016;4:1–10.

466 18. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of
467 gut microbiome variation. *Science.* 2016;352:560–4.

468 19. Laforest-Lapointe I, Messier C, Kembel SW. Tree phyllosphere bacterial communities: exploring
469 the magnitude of intra- and inter-individual variation among host species. *PeerJ.* 2016;4:e2367.

470 20. Bijma P. Multilevel selection 2: estimating the genetic parameters determining inheritance and
471 response to selection. *Genetics.* 2007;175:289-299.

472 21. Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA. A simple and
473 distinctive microbiota associated with honey bees and bumble bees. *Mol Ecol.* 2011;20:619–28.

474 22. Koch H, Schmid-Hempel P. Socially transmitted gut microbiota protect bumble bees against an
475 intestinal parasite. *Proc Natl Acad Sci.* 2011;108:19288–92.

476 23. Eubank S, Guclu H, Kumar VS a, Marathe M V., Srinivasan A, Toroczka Z, et al. Modeling

- 477 Disease Outbreaks in Realistic Urban Social Networks †. *Nonlinear Stud.* 2004;180:6988–6988.
- 478 24. Read JM, Eames KTD, Edmunds WJ. Dynamic social networks and the implications for the spread
479 of infectious disease. *J R Soc Interface.* 2008;5:1001–7.
- 480 25. Naug D. Structure of the social network and its influence on transmission dynamics in a honeybee
481 colony. *Behav Ecol Sociobiol.* 2008;62:1719–25.
- 482 26. Krause J, Croft DP, James R. Social network theory in the behavioural sciences: Potential
483 applications. *Behav Ecol Sociobiol.* 2007;62:15–27.
- 484 27. Archie EA, Tung J. Social behavior and the microbiome. *Curr Opin Behav Sci.* 2015;6:28–34.
- 485 28. Ren T, Boutin S, Humphries MM, Dantzer B, Gorrell JC, Coltman DW, et al. Seasonal , spatial ,
486 and maternal effects on gut microbiome in wild red squirrels. *Microbiome.* 2017;5:1–14.
- 487 29. Griffiths SM, Harrison XA, Weldon C, Wood MD, Pretorius A, Hopkins K, et al. Genetic variability
488 and ontogeny predict microbiome structure in a disease-challenged montane amphibian. *ISME J.*
489 2018;12:2506–2517.
- 490 30. Yuan ML, Dean SH, Longo A V., Rothermel BB, Tuberville TD, Zamudio KR. Kinship, inbreeding
491 and fine-scale spatial structure influence gut microbiota in a hindgut-fermenting tortoise. *Mol Ecol.*
492 2015;24:2521–36.
- 493 31. Lombardo MP. Access to mutualistic endosymbiotic microbes: An underappreciated benefit of
494 group living. *Behav Ecol Sociobiol.* 2008;62:479–97.
- 495 32. Tung J, Barreiro LB, Burns MB, Grenier JC, Lynch J, Grieneisen LE, et al. Social networks predict
496 gut microbiome composition in wild baboons. *Elife.* 2015;2015:1–18.
- 497 33. Grieneisen LE, Livermore J, Alberts S, Tung J, Archie EA. Group living and male dispersal predict
498 the core gut microbiome in wild baboons. *Integr Comp Biol.* 2017;57:770–85.
- 499 34. Moeller AH, Foerster S, Wilson ML, Pusey AE, Hahn BH, Ochman H. Social behavior shapes the
500 chimpanzee pan-microbiome. *Sci Adv.* 2016;2.
- 501 35. Meehan CL, Mcguire MK, Lackey KA, Hagen EH, Williams JE, Roulette J, et al. Social networks ,
502 cooperative breeding , and the human milk microbiome. 2018;1–16.
- 503 36. Lize A, McKay R, Lewis Z. Gut microbiota and kin recognition. *Trends Ecol Evol.* 2013;28:325–6.

- 504 37. Julliand V, Grimm P. The Impact of Diet on the Hindgut Microbiome. *J Equine Vet Sci.* Elsevier
505 Inc.; 2017;52:23–8.
- 506 38. Winton CL, Hegarty MJ, McMahon R, Slavov GT, Mcewan NR, Davies-Morel MCG, et al. Genetic
507 diversity and phylogenetic analysis of native mountain ponies of Britain and Ireland reveals a novel
508 rare population. *Ecol Evol.* 2013;3:934–47.
- 509 39. Stanley CR, Shultz S. Mummy’s boys: Sex differential maternal-offspring bonds in semi-feral
510 horses. *Behaviour.* 2012;149:251–74.
- 511 40. Stanley CR, Mettke-Hofmann C, Hager R, Shultz S. Social stability in semiferal ponies: networks
512 show interannual stability alongside seasonal flexibility. *Anim Behav.* 2018;136:175–84.
- 513 41. Kahle D, Wickham H. ggmap : Spatial Visualization with ggplot2. *R J.* 2013;5:144–61.
- 514 42. RStudio Team. RStudio: Integrated Development for R. 2016. p. RStudio, Inc., Boston, MA URL
515 <http://www.rstudio.c>.
- 516 43. R Core Team. R: A language and environment for statistical computing. R Foundation for
517 Statistical Computing. 2017. p. Vienna, Austria. URL <https://www.R-project.org/>.
- 518 44. Farine DR, Whitehead H. Constructing, conducting and interpreting animal social network
519 analysis. *J Anim Ecol.* 2015;84:1144–63.
- 520 45. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index
521 sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina
522 sequencing platform. *Appl Environ Microbiol.* 2013;79:5112–20.
- 523 46. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-
524 resolution sample inference from Illumina amplicon data. *Nat Methods.* 2016;13:581–3.
- 525 47. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA
526 gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.*
527 2013;41:590–6.
- 528 48. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and “all-species
529 Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res.* 2014;42:643–8.
- 530 49. McMurdie PJ, Holmes S. Phyloseq: An R Package for Reproducible Interactive Analysis and

531 Graphics of Microbiome Census Data. *PLoS One*. 2013;8.

532 50. Oksanen J, Blanchet B, Friendly M, Kindt R, Legendre P, McGlinn D, et al. *vegan: Community*
533 *Ecology Package*. 2018.

534 51. Lahti Leo, Shetty S. Tools for microbiome analysis in R. *Microbiome package version 1.1.10013*.
535 2017. p. <http://microbiome.github.com/microbiome>.

536 52. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York; 2009.

537 53. Cáceres M De, Legendre P. Associations between species and groups of sites: indices and
538 statistical inference. *Ecology*. 2009;90:3566–3574.

539 54. Csárdi G, Nepusz T. The igraph software package for complex network research. *InterJournal*
540 *Complex Syst*. 2006;1695:1695.

541 55. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human
542 gut microbiome viewed across age and geography. *Nature*. 2012;486:222–7.

543 56. Wang J, Linnenbrink M, Künzel S, Fernandes R, Nadeau M-J, Rosenstiel P, et al. Dietary history
544 contributes to enterotype-like clustering and functional metagenomic content in the intestinal
545 microbiome of wild mice. *Proc Natl Acad Sci*. 2014;111:E2703–10.

546 57. Carmody RN, Gerber GK, Luevano JM, Gatti DM, Somes L, Svenson KL, et al. Diet dominates
547 host genotype in shaping the murine gut microbiota. *Cell Host Microbe*. 2015;17:72–84.

548 58. Ben-Yosef M, Behar A, Jurkevitch E, Yuval B. Bacteria-diet interactions affect longevity in the
549 medfly - *Ceratitis capitata*. *J Appl Entomol*. 2008;132:690–4.

550 59. Antwis RE, Haworth RL, Engelmoer DJP, Ogilvy V, Fidgett AL, Preziosi RF. Ex situ diet influences
551 the bacterial community associated with the skin of red-eyed tree frogs (*Agalychnis callidryas*). *PLoS*
552 *One*. 2014;9:e85563.

553 60. Gavriel S, Jurkevitch E, Gazit Y, Yuval B. Bacterially enriched diet improves sexual performance of
554 sterile male Mediterranean fruit flies. *J Appl Entomol*. 2011;135:564–73.

555 61. Lea, JMD, Kerley, GIH, Hrabar, H, Barry, TJ S. Recognition and management of ecological
556 refugees: A case study of the Cape mountain zebra. *Biol Cons*. 2016;203:207-215.

557 62. Ransom JI, Cade BS. *Quantifying Equid Behavior — A Research Ethogram for Free-Roaming*

558 Feral Horses. *Geological Survey Techniques and Methods*. 2009;2-A9:23.

559 63. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, et al. Delivery
560 mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in
561 newborns. *Proc Natl Acad Sci*. 2010;107:11971–5.

562 64. Duranti S, Lugli GA, Mancabelli L, Armanini F, Turrone F, James K, et al. Maternal inheritance of
563 bifidobacterial communities and bifidophages in infants through vertical transmission. *Microbiome*.
564 2017;5:1–13.

565 65. Prest ATL, Kimball AK, Kueneman JG. Host associated bacterial community succession during
566 amphibian development. *Mol Ecol*. 2018;0–2.

567 66. Meale SJ, Li S, Azevedo P, Derakhshani H, Plaizier JC, Khafipour E, et al. Development of
568 ruminal and fecal microbiomes are affected by weaning but not weaning strategy in dairy calves. *Front*
569 *Microbiol*. 2016;7:1–16.

570 67. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. Succession of microbial
571 consortia in the developing infant gut microbiome. *Proc Natl Acad Sci*. 2011;108:4578–85.

572 68. Frese SA, Parker K, Calvert CC, Mills DA. Diet shapes the gut microbiome of pigs during nursing
573 and weaning. *Microbiome*. 2015;3:28.

574 69. Jašarević E, Morrison KE, Bale TL. Sex differences in the gut microbiome–brain axis across the
575 lifespan. *Philos Trans R Soc B Biol Sci*. 2016;371:20150122.

576

577

578

579

580

581

582

583

584 **FIGURE LEGENDS AND TABLE TITLES**

585

586 **Figure 1**

587 NMDS plot of the total microbiome of individual ponies in the study. Larger filled circles indicate the
588 centroid for each individual.

589

590 **Figure 2**

591 Average (\pm SE) Jensen-Shannon divergence metrics for pony microbiome composition (a) within and
592 between individuals; (b) within and between bands; (c) between mothers and their offspring, as well as
593 between juveniles and non-maternal mares and foal-less mares; and (d) between the band stallion
594 and band mares and between all mares within a band. Significantly different results are indicated by *.

595

596 **Figure 3**

597 NMDS plots of the total microbiome of ponies plotted according to band membership (a) and life-stage
598 (b). Larger filled circles indicate group centroids.

599

600 **Figure 4**

601 (a) Map showing spatial distribution of pony bands encountered during sampling. (b) Social network of
602 the sampled individuals with edge width proportional to tie strength between individuals and (c)
603 network visualisation of microbiome distance between individuals.

604

605 **Table 1**

606 Demographic data for each band used in this study.

607

608 **Table 2**

609 Statistical outputs for microbiome adonis analyses.