Defining the resilience of the human salivary microbiota by a 520 days longitudinal study in confined environment: the Mars500 mission

Giovanni Bacci[1], Alessio Mengoni[1], Giovanni Emiliani[2], Carolina Chiellini[3], Edoardo Giovanni Cipriani[1], Giovanna Bianconi[4], Francesco Canganella[4,5]*, Renato Fani[1]*

[1]Department of Biology, University of Florence, Via Madonna del Piano 6, I-50019 Sesto Fiorentino, Italy
[2]Istituto per la Protezione Sostenibile delle Piante, Consiglio Nazionale delle Ricerche, via Madonna del Piano 10, I-50019 Sesto Fiorentino, Italy
[3]Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, I-56124 Pisa, Italy
[4]Department of Biological, Agricultural and Forestry Sciences, Università della Tuscia, Via San Camillo de Lellis snc, I-01100 Viterbo, Italy
[5]Embassy of Italy, 98 Hannam-daero, Hannam-dong, Yongsan-gu, Seoul, South Korea

*Correspondence: renato.fani@unifi.it, canganella@unitus.it

Abstract

The human microbiota plays several roles in health and disease but is often difficult to determine which part is in intimate relationships with the host *vs.* the occasional presence. During the Mars500 mission, six crewmembers lived completely isolated from the outer world for 520 days following standardized diet regimes. The mission constitutes the first spaceflight simulation to Mars and was a unique experiment to determine, in a longitudinal study design, the composition and importance of the resident *vs.* a more variable microbiota—the fraction of the human microbiota that changes in time and according to environmental conditions—in humans. Here we report the characterization of the salivary microbiota from 88 samples taken during and after the mission for a total of 720 days. Amplicon sequencing of the V3-V4 region of 16S rRNA gene was performed and results were analyzed monitoring the diversity of the microbiota while evaluating the effect of the three main variables present in the experimental system: time, diet, and individuality of each subject. Results, though showing statistically significant effects of all three variables, highlighted a main contribution of salivary microbiota personalized features, that is an

28 individual-based resilience of the microbiota. Such findings open the way to consider salivary
29 microbiota under the light of a pronounced personalization even after sharing the same physical
30 space for more than a year.

31 Introduction

32 The host-associated microbiota is stirring the attention from many fields of life science, including
33 basic biology, evolutionary studies, biomedicine, and biotechnology. It is now well known that it
34 plays several roles in modulating the host health and that changes in the composition of the
35 microbiota in specific human body districts or organs (e.g. skin, gut, vagina, lung) may influence
36 the correct functionality of other organs.[1] The concept of holobiont reflects the intimate
37 relationships between the host and the microbiota[2,3] but it is often difficult to determine which
38 part of the host-associated microbiota is in intimate relationships with the host *vs.* an occasional
39 presence. Cross-sectional studies have been used to decipher the more stable, core, microbiota,
40 present in all individuals analyzed, in comparison with the fraction which is more variable,
41 i.e. present in few individuals only (see for instance[4]) and longitudinal analyses helped to
42 understand the temporal stability of the microbiota.[5]

43 The human microbiota is not a single entity but it may have different characteristics and roles.
44 The gut microbiota is expected to be more stable over time than other cavities that are more
45 exposed to the environment—the oral cavity represents one of the first entry point of our body
46 and is thus massively influenced by environmental conditions. The salivary microbiota is known
47 to be affected by both biotic and abiotic factors,[6,7] including age, saliva chemical composition,
48 tongue, and teeth.[8] Consequently, it is still under debate how much of the oral microbiota is
49 stable over time and if this stability can be considered as a tight association with the host.[7,9,10]

50 Given its sensibility to external perturbations, the salivary microbiota could be a good model to
51 inspect the temporal dynamics and subject-by-subject variations impacting the human
52 microbiota, but this sensibility could be a double-edged sword. Even if the disclosure of salivary
53 microbiota temporal stability, and/or subject individuality, could indeed impact on scientific
54 fields spanning from personalized medicine to forensic microbiology, controlling environmental
55 exposures of salivary microbiota is difficult especially during our every day life. Standardize
56 these perturbations implies isolation procedures that are difficult to put in place.

57 Mars500 was the first long-term international study into interplanetary space flights. Managed
58 by the European Space Agency and the Russian Space Agency, it was conducted in 2010-2011
59 when six male volunteers were kept for 520 days in a common confined environment established
60 by the Institute of Biomedical Problems (IBMP) in Moscow, simulating a space flight to Mars.
61 Data from Mars500 mission were studied from various point of view, including behavior,[11] effect
62 of cultural background,[12] cognitive performances,[13] circadian rithms,[14] hormone levels,[15] and
63 surface and gut microbiota.[16,17] Mars500 hence constitutes a unique experiment to determine, in
64 a longitudinal study design, the composition and importance of the resident microbiota vs. a
65 more variable microbiota (changing with time and environmental conditions) in humans. Aim of
66 the work was to inspect the temporal dynamics of salivary microbiota, assessing the effect of diet
67 regimes and individuality, using Mars500 as a unique long-term experiment where subjects were
68 all confined in the same shared environment.

69 # Results

70 ## Salivary microbiota composition during the study

71 To inspect how the salivary microbiota reacts in a confined environment, we characterized
72 samples collected during the entire duration of the Mars500 mission (720 days in total) by 16S
73 rRNA gene amplicon sequencing of the variable region V3-V4. Table S1 summarizes all phases of
74 the mission whereas Figure S1 reports the sampling scheme used in this work. Amplified
75 sequences formed 1890 amplicon sequence variants (ASVs) with a median number of 172.00
76 ASVs per sample (ranging from 81 to 317). A total of 4,337,540 sequences specifically aligned to
77 an ASV resulting in a sequencing depth ranging from 20,084 to 116,809 and a median value of
78 47,044 (for additional information about sequence analysis pipeline and the number of sequence
79 obtained in each pre-processing step see Supplementary material, Supplementary Table S2, and
80 Supplementary Figure S2). All replicates reported an accuracy higher than 0.96 with a
81 Spearman's rank correlation (ρ) that ranged between 0.94 and 0.98 (Supplementary Table S3, S4
82 and Supplementary Figure S3). Rarefaction curves reached a plateau above 15k reads suggesting
83 an adequate sequencing depth for all samples (Supplementary Figure S4). Good's coverage
84 estimator ranged between 99.99% and 100.00% across all samples indicating that roughly 0.01%
85 of the reads in a sample came from ASVs that appear only once in that sample (Supplementary
86 Table S5).

87 Roughly, 99% of sequences aligned to variants that came from known bacterial taxa (Table S2).
88 Supplementary Table S6 shows the overall taxonomic composition of samples whereas Figure S5
89 reports the phylogenetic tree reconstructed from ASVs. At phylum level Firmicutes,
90 Bacteroidetes, Actinobacteria, Proteobacteria, and Fusobacteria accounted for more than 97% of
91 the total number of reads assigned to taxonomically annotated ASVs (Figure 1a and

92 Supplementary Table S6). The total bacterial diversity (namely the alpha diversity) remained
93 constant during the mission with no significant differences detected between the isolation period
94 and the follow-up, across different diets, and across subjects (Table S7 and Figure 1 panels b, c,
95 and d). Also time did not impact bacterial diversity as showed in Figure 1 (random mixed model
96 fitted using crewmembers as random intercept: slope lower than 0.002; Supplementary Table
97 S8).

98 **Effect of food and time on salivary microbiota**

99 We inspected differences across samples (namely beta diversity) using non-metric
100 multidimensional scaling (nMDS) on quantitative and qualitative indexes. Samples showed a
101 similar distribution with all index tested (Figure 2a): Sorensen index and unweighted unifrac
102 distance (qualitative analysis), and Bray-Curtis and weighted unifrac distance (quantitative
103 analysis). As opposed to alpha diversity, subjects, diets, and time significantly contributed to
104 shape the salivary microbiota with different percentage of variance explained depending on the
105 index but never exceeding 10% of the total variance (Figure 2b). For all diversity indexes (except
106 for the Sorensen index which reported a significant effect of subjects) the dispersion of tested
107 factors was homogeneous, meaning that only the composition of samples varied among groups as
108 highlighted by the permutational analysis of variance reported above. Diet impacted on bacterial
109 genera usually present in the salivary microbiota of healty subjects—such as *Actinomyces*,
110 *Veilonella*, and *Fusobacterium*[18]—but also on *Peptostreptococcus*, *Haemophilus*, *Megasphaera*, and
111 *Prevotella*, which have been correlated to different disorders of the oral cavity (such as
112 periodontitis, dental caries, and oral lichen planus).[19–21] Bacterial species classified as
113 *Alloprevotella*, *Fusobacterium*, *Dialister*, *Veilonella*, and *Megasphaera* followed the same pattern:
114 decreasing abundance passing from the the first to the third diet and then back to starting values

115 during the follow up (even if the shift was not significant). Other species like *Haemophilus* and
116 *Prevotella* reported a significantly higher abundance during a single diet—namely *Haemophilus*
117 was more abundant during the follow up and the abundance of *Prevotella* was higher during the
118 first diet. At phylum level, diets impacted more on Bacteroidetes and Firmicutes. Whithin them, 4
119 out of 28 (14%), and 3 out of 7 (43%) genera reported (at least) a significant difference during
120 diet changes (Figure S6).

121 To explore the effect of time on bacterial diversity we used change-point analysis on both within-
122 subject (Figure S7 panel a) and between-subjects diversity (Figure S7 panel b). Within-subject
123 diversity measures changes in the salivary microbiota of each crewmember through time,
124 whereas between-subjects diversity compares the salivary microbiota of different crewmembers
125 at each time point (Figure 3a and b). Three segments significantly divided within-subject
126 diversity with two change-points at 123 days and 480 days. Between-subjects diversity was not
127 segmented since the overall model gave better results than the segmented one according to the
128 genetic algorithm used during optimization (Figure 3b). The overall between-subjects model had
129 an effect size of -0.00004 which means that after 520 days of isolation the overall diversity
130 decreased by 0.02179. The effect of time on within-subject diversity was indeed higher than the
131 one observed for between-subjects diversity. During the first 123 days the effect modeled was
132 0.00103 reflecting an average increase of 0.12636 for all crewmembers. After the first change-
133 point, within-subject diversity started to decrease with a regression parameter of -0.00064
134 (average decrease during the second segment of -0.22835). After the second change point, which
135 roughly matched the end of the isolation period (Figure 3c and d), the within-samples diversity
136 started to increase again. At the end of the follow-up period diversity increased again of 0.29474
137 exceeding the average value detected in the first day of isolation.

138 **Resilience of salivary microbiota**

139 The average abundance of ASVs correlates with their persistence, the number of subjects in
140 which a given ASV was detected at each time point. Figure 4a shows the increasing trend of log-
141 transformed abundance with an R-squared value of 0.72 (b = 10.09, 95% CI [9.81, 10.38]). Time-
142 resolved clustering produced two groups of ASVs: one, called inconsistent micriobiome (Cluster
143 1), included variants detected in a small number of subject at each time point, whereas the other
144 (Cluster 2), called stable microbiota, included variants detected in the vast majority of subjects
145 during the whole mission (Figure 4a and Figure S8, panel a and b). The inconsistent microbiota
146 showed low average persistence in respect with the stable microbiota but it contained the largest
147 amount of variants (1746 ASVs against 144 of stable microbiota). Unlike stable ASVs, subjects
148 lost and acquired inconsistent ASVs both during and after the isolation period (Figure S8 panel b
149 and c). Stable ASVs were detected in roughly 30% of all subjects at each time point (26 samples
150 on 88) with sporadic losses and acquisitions (Figure 4a and Figure S9).

151 We represented the acquisition and loss of bacterial species during the whole mission using
152 networks. At each time point we linked subjects to ASVs detected in their salivary microbiota
153 forming a bipartite network structure which reflected the underlying bacterial community
154 structure. The loss and acquisition of bacterial ASVs was shown in supplementary video S1
155 where green squares represent subjects, red circles represent inconsistent microbiota, and light
156 blue circles represent stable microbiota As shown in the video, the topology of the networks did
157 not change in time, but at each time point subjects acquire/release bacterial species from/into
158 the environment, except for stable ASVs which are shared by most crewmembers and thus
159 (almost) always present in central part of the network. The number of new edges formed and
160 destroyed passing from one time point to another slightly decreased in time (mixed effect model

161 95% CI for formed edges [-0.08, -0.02] and destroyed edges [-0.08, -0.01] Table S9). The end of
162 the isolation period significantly increased the average number of formed edges (namely
163 acquired ASVs) of 28 but the trend was still negative (Figure 4b). The number of formed edges
164 was independent from the number of lost edges (Spearman's ρ = -0.04; p-value = 0.750). The
165 salivary microbiota structure did not change both during and after the isolation period reporting
166 a similar network topology even if we exclude lost ASVs (Figure 4c).

167 **Drivers of diversity**

168 To inspect drivers of beta diversity along and after the isolation period, we fitted a linear model
169 for each ASV detected in the salivary microbiota of crewmembers. The Mars500 mission time-
170 scale was divided into three stages according to the changepoints detected for within-subject
171 diversity. Crewmembers showed a different number of ASVs reporting a trend of diversity
172 similar to the one reported in Figure 3c. The number of ASVs showing a significant effect of time
173 and changepoints ranged from 3 to 28 depending on the subject (Figure 5a and Table S10). Stable
174 ASVs—namely those grouped into Cluster 2 according to their prevalence in time—that showed a
175 significant trend of diversity enriched the saliva of four out of six crewmembers, if compared
176 with the overall occurrence (Figure 5b). The fraction of stable ASVs in subjects 5004 and 5005
177 was roughly six times higher than the average fraction of stable ASVs, whereas subjects 5002 and
178 5006 reported a fraction three times higher than the average.

179 **Discussion**

180 The reported experimental results under Mars500 mission evaluated the temporal dynamics of
181 the human salivary microbiota in a controlled and confined environment. All samples from the 6
182 (male) crewmembers included Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, and

183 Fusobacteria as main phyla and conserved their taxonomic composition along time and among
184 individuals, as previously reported by other studies (see for instance[7,9,22,23]). Despite this
185 conservation, external factors significantly influenced the salivary microbiota composition but
186 their influence was restricted to a low percentage of the community (less than 10% of total
187 variance explained). We found that a great number of species intermittently passed through the
188 salivary microbiota but they neither affected its overall structure nor its taxonomic composition.
189 Despite their number, transient species struggle to thrive in human saliva reporting a low
190 abundance during the whole experiment. On the other hand, a hundred bacterial taxa dominated
191 the salivary microbiota composition of all crewmembers rarely changing host and abundance
192 even when isolation period ended. The isolation time affected bacterial diversity of single
193 individuals but it did not alter the microbiota of the whole crew. The bacterial diversity of
194 crewmembers decreased—if compared within consecutive time points of the same individual—
195 but the effect ended immediately after the isolation period when it started to increase again.
196 External perturbations, impossible to control outside the isolation facility, modulated the salivary
197 microbiota composition when crewmemebrs got out the isolation facility and likely started eating
198 different types of food, getting in touch with other people, or simply visiting different places.

199 The finding obtained under Mars500 mission suggests that sharing the same confined
200 environment—and possibly following the same diet regime with few variation—imbalanced the
201 ecology of the salivary microbiota while reducing its complexity. Unfortunately, we do not know
202 if this effect could—positively or negatively—affect the health of the hosts. However, a loss of
203 strains in the microbiota corresponds to a loss of putative members that could be relevant under
204 changing environmental conditions.[3,4,24] Extending this concept we could suggest that a
205 depauperate microbiota is less reactive to sudden changes in external conditions weaken the

206 host adaptive capacity. However, between-subjects differences remained unaltered suggesting
207 that individuals follow somewhat independent dynamics of their salivary microbiota
208 (i.e. personalized dynamic). The same evidence indicates that crewmembers, though sharing the
209 same environment, did not exchange their salivary microbiota, leading to hypothesize that quite
210 stable personal salivary microbiota features are present in humans and confirming the between-
211 subject effect. The fecal microbiota of the same crewmembers showed an increasing trend of
212 similarity among subjects, especially in relation to a sharing of rare taxa,[17] indicating that salivary
213 microbiota has a more pronounced personalization than fecal microbiota. Indeed, from our data
214 there are no evidences of an ecological succession, as those shown in the fecal microbiota of the
215 same crewmembers. In support of a more pronounced personalized dynamics of salivary
216 microbiota than of fecal microbiota in Mars500 mission, is the finding that the ASVs contributing
217 to the temporal dynamics in the subjects were highly variable (from 3 to 28) indicating that each
218 crewmember salivary microbiota changes were driven by a peculiar (viz. personal) set of
219 microbial taxa.

220 Even if factors such as diet and time influenced the salivary microbiota composition of
221 crewmembers, most details of the Mars500 missions are unknown. The mission was a military
222 experiment and several outcomes are still sealed. The composition of diets for example is
223 unknown and thus any speculation on the effect of particular food intake would not be grounded.
224 Despite these limitations the use of 16S rRNA gene amplicon sequencing, allowed us to detect key
225 features of human salivary microbiota under a condition that is almost impossible to replicate.
226 The complete isolation of the participants of the mission made possible the first observation of
227 salivary microbiota composition minimizing the effect of external perturbations.

228 In conclusion, the reported longitudinal analysis of human salivary microbiota confirmed the
229 stability of the microbiota over time and suggested the presence of resilient personalized
230 taxonomic features, which may deserve further attention. This study allowed clearly to
231 determine the contribution of a stable and confined environment, as that of Mars500 mission, in
232 reducing the microbiota diversity and to show the effect of a controlled diet on salivary
233 microbiota.

234 Methods

235 The Mars500 experiment

236 Mars500 mission was conducted in 2010-2011 by Russia's Institute of Biomedical Problems
237 (IBMP), with extensive participation by the European Space Agency (ESA) as part of the
238 European Programme for Life and Physical Sciences (ELIPS) to prepare for future human
239 missions to the Moon and Mars. The whole project consisted of three isolation studies: a 14-day
240 pilot study to test facilities and procedures used during the simulation, a 105-day pilot study
241 involving six crewmembers, and a 520-day study that simulated a complete space flight to Mars
242 and back. Mars500 crew was composed of six male volunteers. All crewmembers were confined
243 in the same living space from the 3rd of June 2010 till the 4 of November 2011 when they finally
244 stepped out of the isolation facility to come back to their normal activities. During the mission the
245 crew was hermetically isolated from the rest of the IBMP facility. Crewmember received three
246 type of diets, a so called "first variant" (FV), "third variant" (TV) and, after the experiment,
247 returned to a normal (non supervised) diet regime (NR). Detailed information on the experiment
248 are reported in Supplemental Information file. Further details are also reported in the companion
249 Mars500 microbiology paper.[17]

250 **Collecting salivary samples**

251 Saliva samples were collected individually, based on the scientific protocol and pre-confinement
252 training, with 5ml sterile vials (Nalgene V5257-250EA). Upon completion of the saliva sampling
253 all samples from one sampling event were put into the hatch. After that, they were removed by
254 the responsible person of the IBMP and stored at -80°C. After being stored at -80°C in the
255 laboratories of the IBMP for periods of at least 4 days up to 6 months, the samples were sent via
256 World Courier. Shipping from Moscow to the University of Tuscia, Viterbo – Italy. The shipment
257 was performed in three batches on dry ice to avoid repeated freeze-thaw cycles which lead to
258 reduction of microbial viability. Upon arrival, samples were kept at -80°C until processing.
259 Salivary samples were collected by crewmembers during and after the permanence in the
260 isolation facility (Table 1): 42 samples were collected during the first simulated journey from
261 Earth to Mars (seven time-points), 30 samples were collected during the simulated trip back
262 home from Mars to Earth (five time-points), and other 16 samples were collected when
263 crewmembers came back to their normal activities and were followed for additional 200 days
264 (three time-points). Unfortunately, two samples collected in the latter stage gave no good quality
265 DNA and were thus discharged. Crewmembers did not collect salivary samples during the
266 simulated landing on Mars where three of them—which simulated the landing on a separate
267 module—used a different food variant. For this reason the second food variant used was not
268 reported in the work passing from the first food variant (FV) directly to the third food variant
269 (TV).

270 **Sequencing of salivary samples**

271 DNA was extracted from salivary samples stored at -80°C using a conventional bead-beating
272 protocol (DNeasy PowerSoil Kit, Mobio). After fluorimetry quantification (Qubit), 20 ng of

273 environmental DNA were used as template for amplification of 16S rRNA gene using V3-V4
274 primers (341F and 785R) as previously reported.[25,26] Libraries were constructed and sequenced
275 on a MiSeq apparatus[27] (Illumina) by BMR Genomics (Padua, Italy). Sequences are deposited at
276 ENA database under the accession ERP119217.

277 **Amplicon sequence variant reconstruction**

278 The DADA2 pipeline (version 1.14.1) was used to reconstruct amplicon sequence variants (ASVs)
279 from illumina reads 28. Both ASV reconstruction and statistical analyses were performed in the R
280 environment version 3.4.3 (http://www.R-project.org). For a complete description of all the step
281 performed see Supplementary material section. Briefly, primers used for V3-V4 amplification
282 were detected and removed using cutadapt version 1.15.[29] Low quality reads were discarded
283 using the `filterAndTrim` function with an expected error threshold of two for both forward and
284 reverse read pairs. Denoising was performed using the `dada` function after error rate modelling
285 (`learnErrors` function). Denoised reads were then merged discarding those with any
286 mismatches and/or an overlap length shorter than 20bp ('mergePairs' function). Chimeric
287 sequences were removed using the `removeBimeraDenovo` function. Taxonomical classification
288 was performed using DECIPHER package version 2.14.0 against the latest version of the pre-
289 formatted Silva small-subunit reference database (SSU version 132 available at:
290 http://www2.decipher.codes/Downloads.html).[30,31] All sequences classified as chloroplasts,
291 mitochondria, Archaea and Eukarya were removed. A summary of retained reads in each step is
292 reported in Table S1 and in Figure S1.

293 **Diversity estimation**

294 Bacterial diversity in each sample was computed using inverse Simpson index as implemented in
295 the `diversity` function of vegan package. Differences according to crewmembers, permanence in

296 the isolation facility, and food variants were inspected using one-way analysis of variance
297 (ANOVA). The effect of time was modeled using linear mixed models with fixed slope and random
298 intercept. Since alpha diversity was measured multiple times on the same statistical units,
299 crewmembers were used as random intercept factor.

300 Diversity across samples was inspected using different approaches. Qualitative and quantitative
301 indexes were used to infer pairwise distances between samples. Qualitative indexes are binary
302 indexes which take into account presence/absence of species to compute distances between
303 samples whereas quantitative indexes are mainly based on the abundance of species.[32] Sorensen
304 index[33] and un-weighted UniFrac distance[34] were used as qualitative indexes whereas Bray-
305 Curtis dissimilarity[35] and weighted UniFrac distance[34] were used as quantitative indexes. UniFrac
306 distances were computed using the `distance` function of the phyloseq R package version 1.30.0[36]
307 whereas Sorensen and Bray-Curtis dissimilarity indexes were computed using the `vegdist`
308 function of the R package vegan verison 2.5-6.[37] Differences between salivary microbiota
309 composition of the same crewmember at consecutive time-points (within-subject diversity) were
310 computed using the `TBI` function of the adespatial R package verison 0.3-8.[38] Packages `vegan` and
311 `adespatial` use the same definition of Sorensen and Bray-Curtis distance.

312 Sorensen index is defined as:

$$\frac{(A + B - 2J)}{(A + B)}$$

313 where A and B are the numbers of ASVs on compared samples, and J is the number of the ASVs
314 shared by both samples.

315 Bray-Curtis index is defined as:

$$b_{ik} = \frac{\sum_{j=1}^{J} |\, n_{ij} - n_{kj} \,|}{\sum_{j=1}^{J} n_{ij} + n_{kj}}$$

where b_{ik} is the Bray-Curtis distance between sample i and k with J number of species, n_{ij} is the abundance of species j in sample i, and n_{kj} is the abundance of species j in sample k[39]. The `adespatial` package computes the index by splitting the total diversity into 3 components: A (the unscaled similarity between two samples), B (the species loss between two samples), and C (the species gain between two samples). The computation is then performed using the formula $(B + C)/(2A + B + C)$ giving the same results as the formalization given above.

Qualitative indexes rely on the assumption that all taxa are equally contributing to bacterial diversity independently from their abundance. For this reason, even extremely rare taxa may be relevant in shaping sample distribution. To relax this assumption ASVs detected in less than 5% of samples (4 samples overall) with an abundance lower than 10 were filtered out before diversity calculation.

Distances across samples were reported using non-metric multidimensional scaling (nMDS) as implemented in the `metaMDS` function of the vegan package, with 300 random starts and monotone regression.[40,41] To test the effect of food variants, crewmembers, and time in shaping the salivary microbiota we used permutational multivariate analysis of variance on distance matrices obtained above (`adonis2` function of the vegan package with 1,000 permutations). The proportion of sum of squares from the total (namely the R^2 value of permutational analysis) was used to report the percentage of variance explained by each factor included in the analysis. Before testing for differences in bacterial composition among groups is advisable to make sure that groups are homogeneously dispersed, otherwise permutational tests (such as `adonis`) may report significant results entirely due to uneven dispersion. To distinguish between actual

337 differences in composition or differences due to dispersion we used the `betadisper` and `anova`
338 functions (`vegan` package). P-values obtained were corrected using Benjamini & Hochberg
339 correction (also known as false discovery rate)[42]. To avoid possible biases induced by uneven
340 sequencing depths, read counts were scaled using DESeq2 before diversity calculation (`counts`
341 function) 43. Scaled counts were additionally transformed using the square root of the Wisconsin
342 double standardized counts (`wisconsin` function of the vegan package).

343 **Influence of time on bacterial diversity**

344 Salivary microbiota may be affected by several factors. Sharing the same environment for a
345 prolonged period of time may alter the composition of salivary microbiota at different levels. The
346 bacterial composition may be altered within the same individual taken at consecutive time points
347 but even between multiple individuals at each time point. To inspect both of these components,
348 bacterial diversity within and between subjects, calculated as reported above, was modeled
349 through time. Change-point analysis was used to identify specific time points which led to a
350 decrease/increase of diversity. The optimal positioning and number of change-points for each
351 crewmember was identified using a non-parametric cost function as implemented in the `cpt.np`
352 function of the changepoint.np R package, version 1.0.1.[44] The pruned exact linear time algorithm
353 (PELT)[45] was used to detect temporal changes in diversity within the same subject and between
354 different subjects. The PELT algorithm searches for an optimal solution by minimizing the cost of
355 different segmentation. We used the modified Bayes information criterion penalty term (MBIC)[46]
356 as penalty function for cost minimization by the algorithm. Since PELT algorithm is exact, a
357 solution is always found for each time series so, to avoid inflation of change-points due to the
358 presence of data coming from six different subjects, a genetic algorithm was used to fine-tune the

359 analysis. All change-points detected were used as starting point of the genetic algorithm and a
360 fitness function was defined as:

$$\frac{1}{\sum_{k=1}^{n} R\,MSE(lmm_k)}$$

361 where, $RMSE(lmm_k)$ stands form the root mean square error of the generalized linear model
362 constructed on the k segment and n is the number of segments defined by change-point analysis.
363 High error corresponds to a low fitness value whereas low error corresponds to a high fitness
364 value. At each step of iteration, the genetic algorithm will keep segmentation that lead to linear
365 models with a low RMSE and discard those leading to high error models. We implemented this
366 algorithm using the R package GA[47] verison 3.2 with a population size of 200. Generalized linear
367 model were fitted using crewmembers as random intercept and p-value were computed with the
368 Satterthwaite's degrees of freedom method as implemented in the package lmerTest[48,49] version
369 3.1-1.

370 We assessed the persistence of the salivary microbiota across subjects using dynamic time
371 warping algorithm implemented in the dtwclust R package (version 5.5.6)[50]. At each time point
372 the number of subjects in which a given ASV was detected (namely the ASV's persistence) was
373 reported together with its abundance. Persistence matrix was scaled and centered before
374 clustering. Centering was performed by subtracting the mean of each ASV from their persistence
375 whereas scaling was performed by dividing each value by its standard deviation. The relation
376 between persistence and abundance was tested by fitting a linear model ('lm' function of R stat
377 package). All detail s about clustering and modelling were reported in Supplementary methods.
378 Persistence across subjects was also used for network construction: at each time point we
379 constructed a bipartite network by linking subjects with ASVs that were present in their salivary

microbiota. Subjects were represented using squared nodes whereas ASVs were represented using round circles colored according to the groups defined above. Doing so, we generated twelve bipartite, acyclic, and undirected networks representing the salivary microbiota of all subjects at different time points. The network R package (version 1.16.0) was used for network reconstruction whereas the package ggnetwork (version 0.5.8) was used for plotting.[51,52] The effect of time and the end of the isolation period on the number of formed/destroyed edges were tested using mixed effect models with random intercept. Subjects were taken as random intercept whereas the time and the end of the isolation period as fixed effects. P-values were computed as discussed above.

Differences along time, for each ASVs detected, were inspected by selecting drivers of within-subject beta diversity. Absolute differences between consecutive time points were fitted using linear models and the effect of time and changepoints was inspected. The slope of the models was used to assess the trend of selected ASVs in each stage. We selected ASVs showing a trend similar to what we found during changepoint detection to focus the analysis only on components of the salivary microbiota of each subject contributing to the total diversity. Finally, To inspect if "drivers of diversity" detected were more present in stable or inconsistent salivary microbiota, a hypergeometric test was performed. The test compared the the occurrence of stable ASVs (Cluster 2) in the overall population against the microbiota of single subjects seraching for significant enrichments in respect with the microbiota distribution in all subjects.[53]

Tables

Table 1: Number of salivary samples collected during the study.

Subject	Earth to Mars (First diet)	Mars to Earth (Third diet)	Follow-up (No diet)	Total

5001	7	5	3	15
5002	7	5	3	15
5003	7	5	3	15
5004	7	5	3	15
5005	7	5	2	14
5006	7	5	2	14
Total	42	30	16	88

The number of samples collected during each step of the study was reported for each crewmember. Marginal totals were added for subjects and simulated journeys together with the grand total that was reported in the bottom right corner of the table.

Table 2: Temporal changes of salivary microbiota.

Days	*b*	SE	*t*	*df*	*p*
Within-subject					
1 - 123	0.00103	0.00027	3.86	12.00	.0023
124 - 480	-0.00064	0.00012	-5.19	36.00	< .0001
481 - 720	0.00123	0.00020	6.16	14.23	< .0001
Between-subjects					
1 - 720	-0.00004	0.00001	-3.89	424.19	.0001

Within-sample diversity was divided into three segments following change-point analysis whereas between-samples diversity was modeled on the full time period since no change-points were detected. Results of mixed effect models fitted for each segment were reported in the table. *b*, regression parameter (slope of the model); SE, standard error; *t*, t-value (also known as "standardized" regression parameter); *df*, degrees of freedom; *p*, p-value.

Figure legend

Figure 1: Salivary microbiota diversity along Mars500 mission and follow-up. a) Distribution of the main bacterial classes. Panels were divided according to crewmembers (vertically) and diets (horizontally). ASVs with a relative cumulative frequency lower than 5% in all samples were collapsed into a single group called "Other". b) Differences in alpha diversity—reported using the inverse Simpson index just like panels c, d, and e—between samples collected during and after the isolation period. c) Differences across diets (FV, first variant; TV, third variant; NR, normal diet). d) Differences among subjects. e) Differences along the whole mission and during the follow-up. Points are the average diversity values among subjects whereas errorbars represent the 95% confidence interval around the mean. The red line represents the population effect of the linear mixed model whose coefficients are reported in Table S8.

Figure 2: Microbial assemblage variation according to diet, crewmembers, and time. a) Non-metric multidimensional scaling based on different beta-diversity indexes (reported on the top of each panel). Samples were colored according to crewmembers whereas the point shape represents the type of diet (FV, first variant; TV, third variant; NR, normal diet). The dispersion of groups was tested for homogeneity and results were reported on the top of each oridnation (a p-value higher than 0.05 means that dispersions are homogeneous). b) Permutational multivariate analysis of variance using distance matrices based on the same indexes reported in panel a. The R^2 values associated with each factor used in the analysis is reported in the horizontal axis whereas asterisks report the significance level of each factor (*, p-value < 0.05; **, p-value < 0.01). Colors represent the different factors modeled in the analysis. For additional information about diets, sampling point and crewmembers see Supplementary information and Supplementary Figure S1.

Figure 3: Crewmembers' salivary microbiota composition in time. a) Bray-Curtis (also known as quantitative Sorensen) index has been used to inspect distances between and within-subject during isolation and follow-up. Between-subjects diversity was computed by comparing the salivary microbiota of each subjects at each timepoint (gray arrows); within-subject diversity was computed by comparing the salivary microbiota of the same subject over time (black arrows). b) Change-point analysis revealed changes in salivary microbiota composition of each subject (CP detection). Genetic algorithm and linear modelling detected increasing/decreasing patterns along time (GA optimization). Finally, we fit a linear mixed-model for each segment detected using crewmembers as random intercept (Modelling). c and d) Results obtained following the pipeline reported in "b" for within- and between-samples differences. Diets were reported in the bottom part of the plots using different colors (FV, fist food variant; SV, second food variant; TV, third food variant). Since crewmemebrs ate freely during the follow-up no diet was reported in the plot.

Figure 4: Bacterial community structure in time. a) Persistence and abundance of ASVs detected. Persistence was expressed as the number of subject in which an ASV was detected whereas abundance was expressed as log-normalized number of reads assigned to that ASV (the black line represent results of the linear fitting: 95% CI [9.81, 10.38], $t(1926) = 70.27$, $p < 0.001$). Cluster 1 (reported in red) was composed of ASVs with a lower persistence and abundance than Cluster 2 (reported in blue). b) Number of edges formed (green) and destroyed (red) at each time point. Lines represent the result of two linear mixed models with subject as random intercept (Table S9). The dashed line represents the effect at the end of the isolation—it was reported only for formed edges since the number of destroyed edges was not significantly impacted by the isolation. c) Community networks at each time point. Days are reported at the top of each network whereas nodes with no edges (namely ASVs not detected at a given time point) were not reported to save space for graphical representation.

Figure 5: Drivers of diversity for each crewmember. a) Single ASVs showing a trend of diversity along time similar to the whole within-subject diversity were reported. Points represent absolute differences between two consecutive time points whereas colors report the segmentation detected through change-point analysis (Figure 3c). Lines show the predicted

differences using linear modelling. We standardize differences to represent ASVs with different ranges of values in the same panel. b) Enrichment analysis of memebers of Cluster 2 in respect with the population. The firs bar on the left represents the overall fraction of AVSs assigned to Cluster 2 whereas the other bars report the fraction of ASVs ssigned to the same cluster for each subject. Adjusted p-values were reported using a single asterisk ($p < 0.05$) or two ($p < 0.01$).

Acknowledgements

The MARS500 Programme was financed by the European Programme for Life and Physical Sciences in Space (ELIPS). The financial support of the Italian Space Agency (contract I/011/11/0) is highly remarked and acknowledged. This work was partially supported by BMR Genomics for sequencing of salivary microbiota. The authors would like to tank Professor Nicola Segata for his precious suggestions on microbiota analyses and comments on the work.

References

1. Lloyd-Price, J., Abu-Ali, G. & Huttenhower, C. The healthy human microbiome. *Genome Medicine* **8**, 51 (2016).

2. Bordenstein, K. R., Seth R. AND Theis. Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLOS Biology* **13**, 1–23 (2015).

3. Rosenberg, E. & Zilber-Rosenberg, I. Microbes drive evolution of animals and plants: The hologenome concept. *mBio* **7**, (2016).

4. Theis, K. R. *et al.* Getting the hologenome concept right: An eco-evolutionary framework for hosts and their microbiomes. *mSystems* **1**, (2016).

5. Clarke, S. F. *et al.* Targeting the microbiota to address diet-induced obesity: A time dependent challenge. *PloS one* **8**, e65790–e65790 (2013).

6. Shaw, L. *et al.* The human salivary microbiome is shaped by shared environment rather than genetics: Evidence from a large family of closely related individuals. *MBio* **8**, e01237–17 (2017).

7. Nasidze, I., Li, J., Quinque, D., Tang, K. & Stoneking, M. Global diversity in the human salivary microbiome. *Genome research* **19**, 636–643 (2009).

8. Xu, X. *et al.* Oral cavity contains distinct niches with dynamic microbial communities. *Environmental microbiology* **17**, 699–710 (2015).

9. Cameron, S. J. S., Huws, S. A., Hegarty, M. J., Smith, D. P. M. & Mur, L. A. J. The human salivary microbiome exhibits temporal stability in bacterial diversity. *FEMS Microbiology Ecology* **91**, (2015).

10. Mukherjee, C., Beall, C. J., Griffen, A. L. & Leys, E. High-resolution isr amplicon sequencing reveals personalized oral microbiome. *Microbiome* **6**, 153 (2018).

11. Tafforin, C. The mars-500 crew in daily life activities: An ethological study. *Acta Astronautica* **91**, 69–76 (2013).

12. Tafforin, C. Time effects, cultural influences, and individual differences in crew behavior during the mars-500 experiment. *Aviation, space, and environmental medicine* **84**, 1082–1086 (2013).

13. Schneider, S. *et al.* The influence of exercise on prefrontal cortex activity and cognitive performance during a simulated space flight to mars (mars500). *Behavioural brain research* **236**, 1–7 (2013).

14. Vigo, D. E. *et al.* Circadian rhythm of autonomic cardiovascular control during mars500 simulated mission to mars. *Aviation, space, and environmental medicine* **84**, 1023–1028 (2013).

15. Wang, Y. *et al.* During the long way to mars: Effects of 520 days of confinement (mars500) on the assessment of affective stimuli and stage alteration in mood and plasma hormone levels. *PloS one* **9**, e87087 (2014).

16. Schwendner, P. *et al.* Preparing for the crewed mars journey: Microbiota dynamics in the confined mars500 habitat during simulated mars flight and landing. *Microbiome* **5**, 129 (2017).

17. Turroni, S. *et al.* Temporal dynamics of the gut microbiota in people sharing a confined environment, a 520-day ground-based space simulation, mars500. *Microbiome* **5**, 39 (2017).

18. Dewhirst, F. E. *et al.* The human oral microbiome. *Journal of bacteriology* **192**, 5002–5017 (2010).

19. Abusleme, L. *et al.* The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *The ISME Journal* **7**, 1016–1025 (2013).

20. Xu, L. *et al.* Dynamic alterations in salivary microbiota related to dental caries and age in preschool children with deciduous dentition: A 2-year follow-up study. *Frontiers in Physiology* **9**, 342 (2018).

21. Wang, K. *et al.* Preliminary analysis of salivary microbiome and their potential roles in oral lichen planus. *Scientific Reports* **6**, 22943 (2016).

22. Belstrøm, P. A. B., Daniel AND Holmstrup. Temporal stability of the salivary microbiota in oral health. *PLOS ONE* **11**, 1–9 (2016).

23. Hall, M. W. *et al.* Inter-personal diversity and temporal dynamics of dental, tongue, and salivary microbiota in the healthy oral cavity. *npj Biofilms and Microbiomes* **3**, 2 (2017).

24. Leggat, W. *et al.* The hologenome theory disregards the coral holobiont. *Nature Reviews Microbiology* **5**, 826–826 (2007).

529 25. Perrin, E. *et al.* Furnishing spaceship environment: Evaluation of bacterial biofilms on
530 different materials used inside international space station. *Research in Microbiology* **169**, 289–
531 295 (2018).

532 26. Bacci, G. *et al.* Applying predictive models to decipher rhizobacterial modifications in
533 common reed die-back affected populations. *Science of The Total Environment* **642**, 708–722
534 (2018).

535 27. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the illumina
536 hiseq and miseq platforms. *The ISME Journal* **6**, 1621–1624 (2012).

537 28. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from illumina amplicon data.
538 *Nature methods* **13**, 581 (2016).

539 29. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
540 *EMBnet. journal* **17**, 10–12 (2011).

541 30. Wright, E. S. Using decipher v2. 0 to analyze big biological sequence data in r. *R Journal* **8**,
542 (2016).

543 31. Quast, C. *et al.* The silva ribosomal rna gene database project: Improved data processing and
544 web-based tools. *Nucleic acids research* **41**, D590–D596 (2012).

545 32. Lozupone, C. A., Hamady, M., Kelley, S. T. & Knight, R. Quantitative and qualitative β diversity
546 measures lead to different insights into factors that structure microbial communities. *Appl.*
547 *Environ. Microbiol.* **73**, 1576–1585 (2007).

548 33. Sørensen, T. J. *A method of establishing groups of equal amplitude in plant sociology based on*
549 *similarity of species content and its application to analyses of the vegetation on danish commons*. (I
550 kommission hos E. Munksgaard, 1948).

551 34. Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J. & Knight, R. UniFrac: An effective
552 distance metric for microbial community comparison. *The ISME journal* **5**, 169 (2011).

553 35. Bray, J. R. & Curtis, J. T. An ordination of the upland forest communities of southern wisconsin.
554 *Ecological Monographs* **27**, 325–349 (1957).

555 36. McMurdie, P. J. & Holmes, S. Phyloseq: An r package for reproducible interactive analysis and
556 graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).

557 37. Oksanen, J. *et al. Vegan: Community ecology package*. (2019).

558 38. Dray, S. *et al. Adespatial: Multivariate multiscale spatial analysis*. (2019).

559 39. Odum, E. P. Bird populations of the highlands (north carolina) plateau in relation to plant
560 succession and avian invasion. *Ecology* **31**, 587–605 (1950).

561 40. Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric
562 hypothesis. *Psychometrika* **29**, 1–27 (1964).

41. Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika* **29**, 115–129 (1964).

42. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289–300 (1995).

43. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology* **15**, 550 (2014).

44. Haynes, K. & Killick, R. *Changepoint.np: Methods for nonparametric changepoint detection*. (2019).

45. Killick, R., Fearnhead, P. & Eckley, I. A. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association* **107**, 1590–1598 (2012).

46. Zhang, N. R. & Siegmund, D. O. A modified bayes information criterion with applications to the analysis of comparative genomic hybridization data. *Biometrics* **63**, 22–32 (2007).

47. Scrucca, L. GA: A package for genetic algorithms in R. *Journal of Statistical Software* **53**, 1–37 (2013).

48. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* **82**, 1–26 (2017).

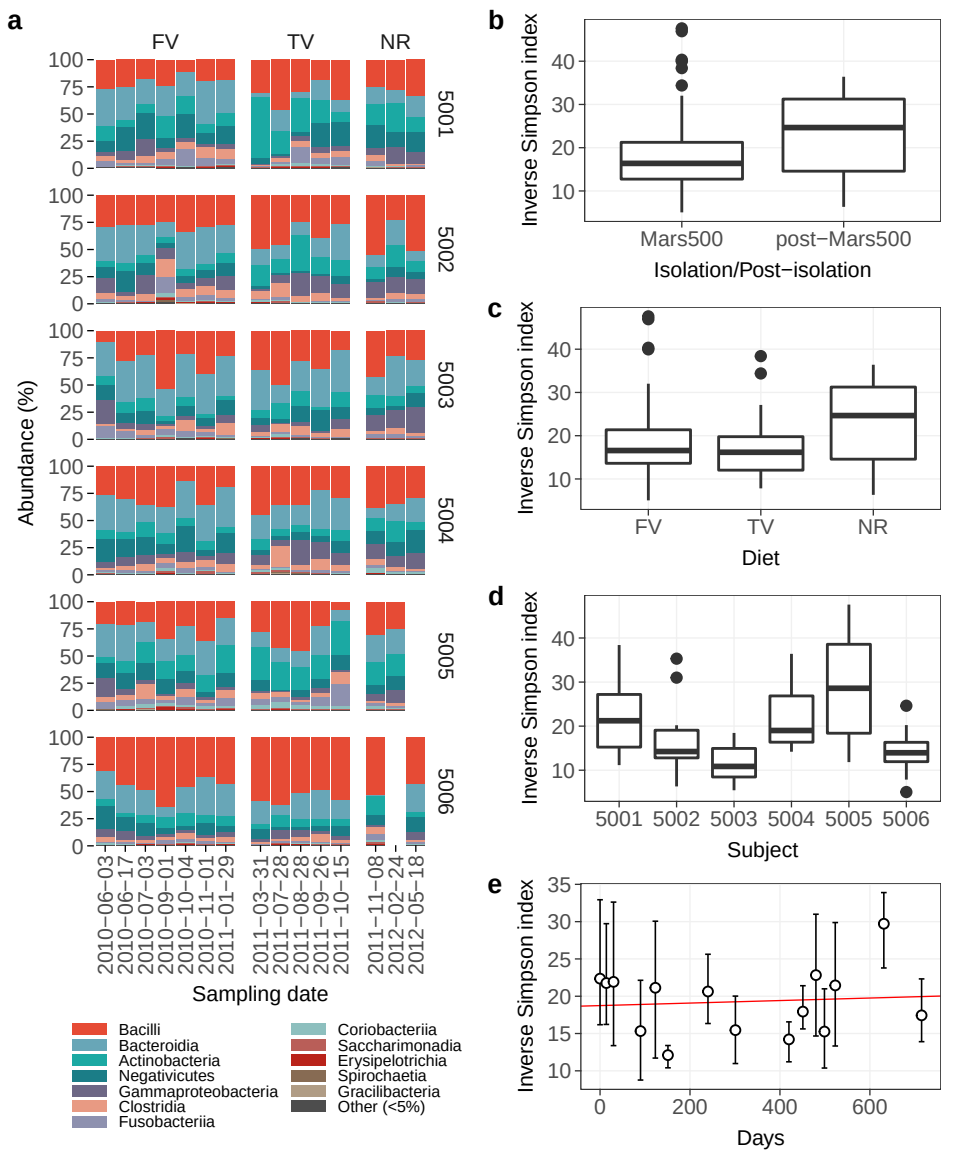
49. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48 (2015).

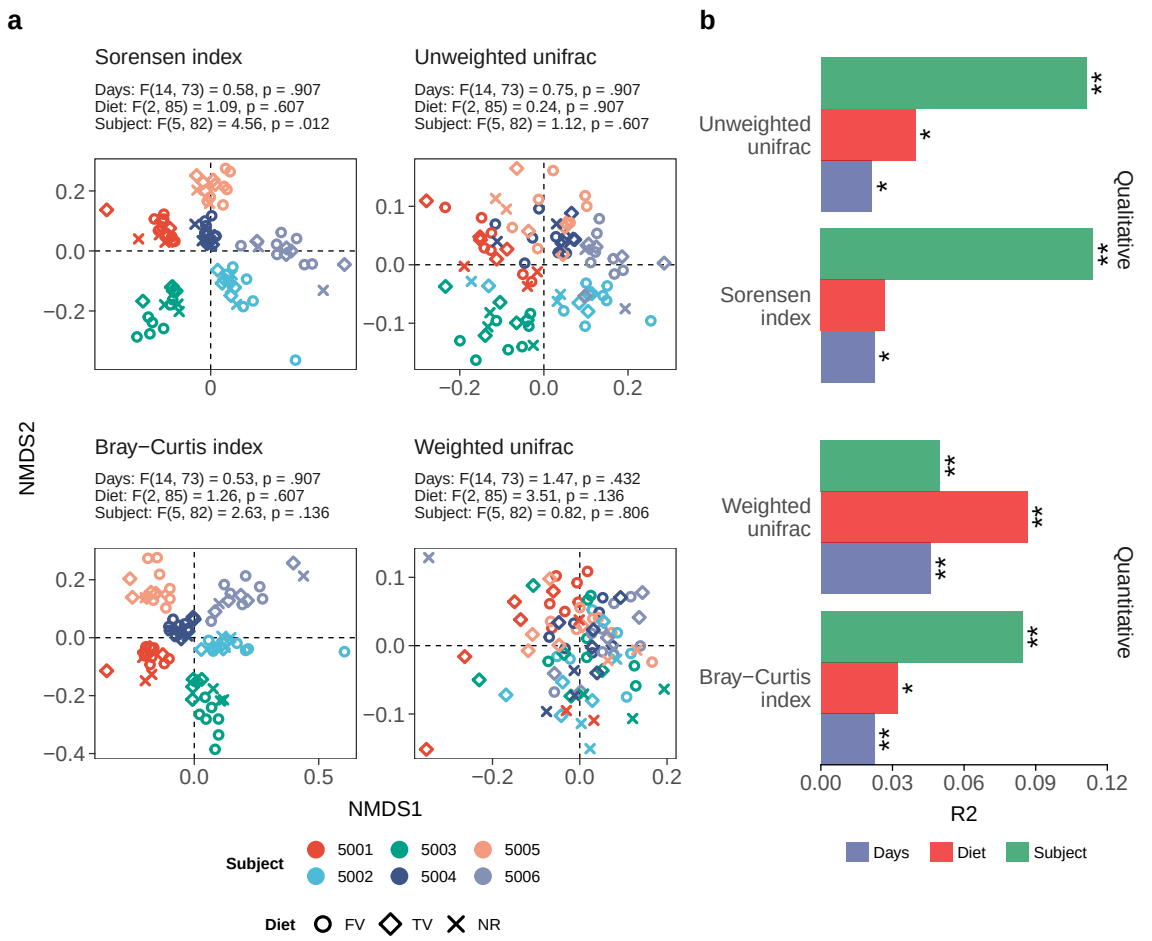
50. Sardá-Espinosa, A. Time-series clustering in r using the dtwclust package. *The R Journal* (2019). doi:10.32614/RJ-2019-023

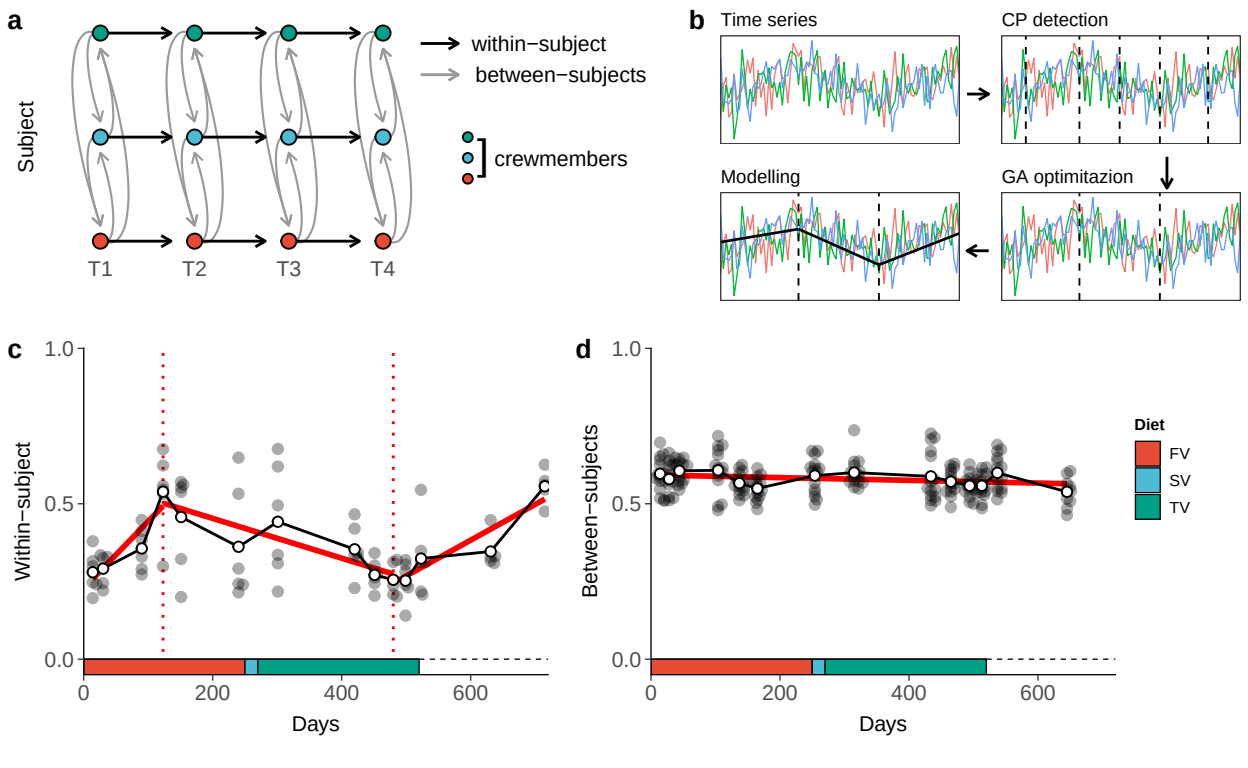
51. Briatte, F. *Ggnetwork: Geometries to plot networks with 'ggplot2'*. (2020).

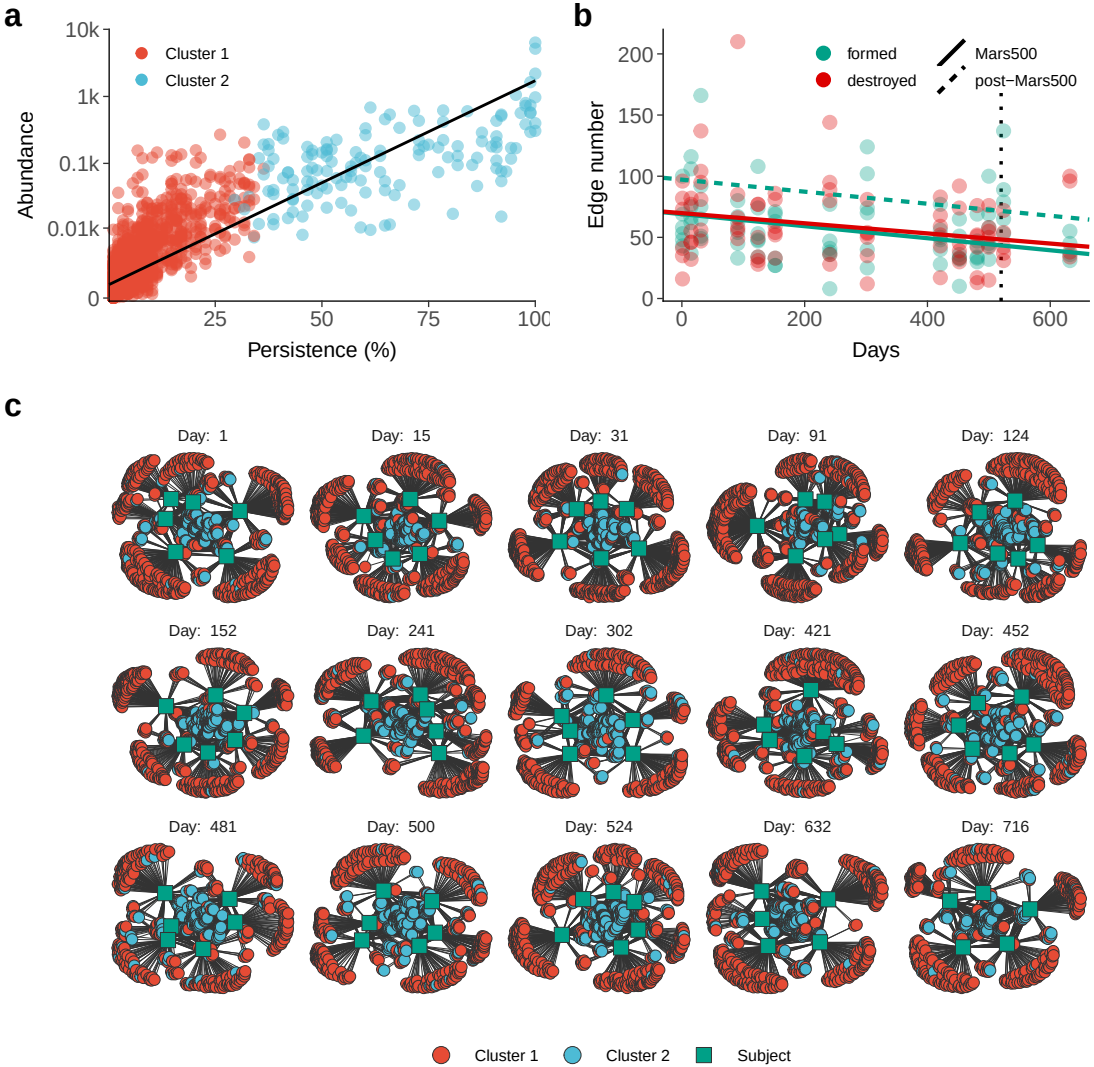
52. Butts, C. T. Network: A package for managing relational data in r. *Journal of Statistical Software* **24**, (2008).

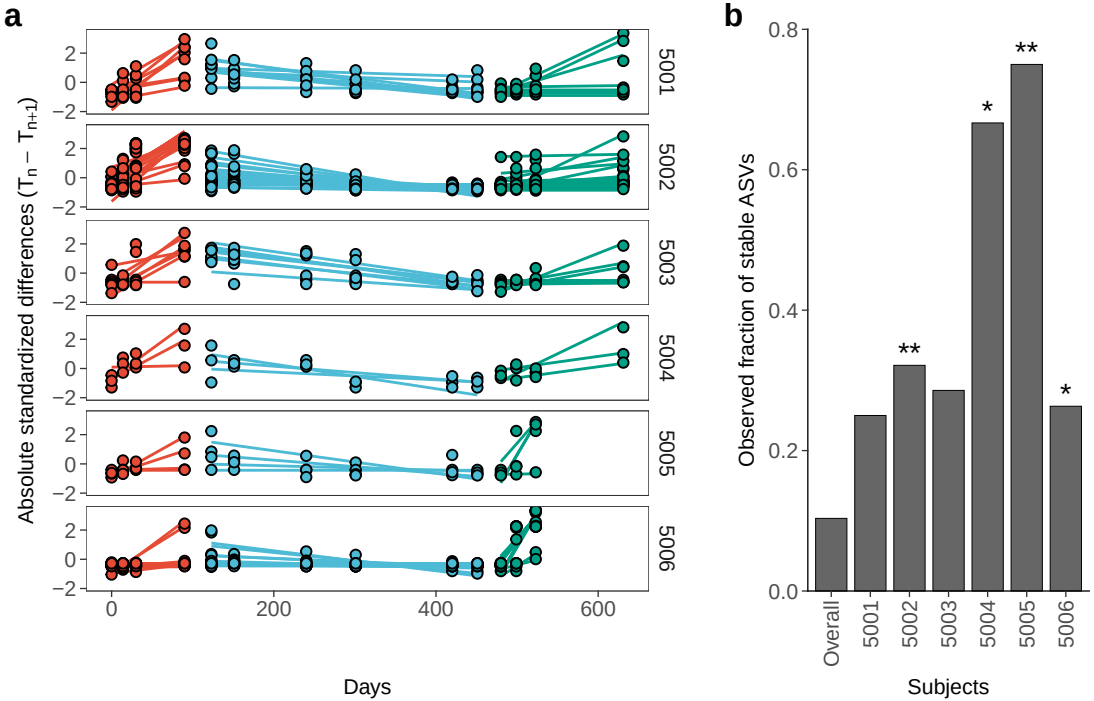
53. Rivals, I., Personnaz, L., Taing, L. & Potier, M.-C. Enrichment or depletion of a go category within a class of genes: Which test? *Bioinformatics* **23**, 401–407 (2006).


a
FV
TV
NR
5001
5002
5003
5004
5005
5006
Abundance (%)
100
75
50
25
0
2010−06−03
2010−06−17
2010−07−03
2010−09−01
2010−10−04
2010−11−01
2011−01−29
2011−03−31
2011−07−28
2011−08−28
2011−09−26
2011−10−15
2011−11−08
2012−02−24
2012−05−18
Sampling date
Bacilli
Bacteroidia
Actinobacteria
Negativicutes
Gammaproteobacteria
Clostridia
Fusobacteriia
Coriobacteriia
Saccharimonadia
Erysipelotrichia
Spirochaetia
Gracilibacteria
Other (<5%)
b
Inverse Simpson index
Mars500
post−Mars500
Isolation/Post−isolation
c
Inverse Simpson index
FV
TV
NR
Diet
d
Inverse Simpson index
5001
5002
5003
5004
5005
5006
Subject
e
Inverse Simpson index
0
200
400
600
Days


a
Sorensen index
Days: F(14, 73) = 0.58, p = .907
Diet: F(2, 85) = 1.09, p = .607
Subject: F(5, 82) = 4.56, p = .012
Unweighted unifrac
Days: F(14, 73) = 0.75, p = .907
Diet: F(2, 85) = 0.24, p = .907
Subject: F(5, 82) = 1.12, p = .607
Bray−Curtis index
Days: F(14, 73) = 0.53, p = .907
Diet: F(2, 85) = 1.26, p = .607
Subject: F(5, 82) = 2.63, p = .136
Weighted unifrac
Days: F(14, 73) = 1.47, p = .432
Diet: F(2, 85) = 3.51, p = .136
Subject: F(5, 82) = 0.82, p = .806
NMDS2
NMDS1
Subject
5001
5002
5003
5004
5005
5006
Diet
FV
TV
NR
b
Unweighted unifrac
Sorensen index
Qualitative
Weighted unifrac
Bray−Curtis index
Quantitative
R2
Days
Diet
Subject


a
Subject
T1
T2
T3
T4
within−subject
between−subjects
crewmembers
b
Time series
CP detection
Modelling
GA optimitazion
c
Within−subject
1.0
0.5
0.0
0
200
400
600
Days
d
Between−subjects
1.0
0.5
0.0
0
200
400
600
Days
Diet
FV
SV
TV


a
Cluster 1
Cluster 2
Abundance
10k
1k
0.1k
0.01k
0
25
50
75
100
Persistence (%)
b
formed
destroyed
Mars500
post−Mars500
Edge number
200
150
100
50
0
0
200
400
600
Days
c
Day: 1
Day: 15
Day: 31
Day: 91
Day: 124
Day: 152
Day: 241
Day: 302
Day: 421
Day: 452
Day: 481
Day: 500
Day: 524
Day: 632
Day: 716
Cluster 1
Cluster 2
Subject


a
Absolute standardized differences ($T_n - T_{n+1}$)
2
0
−2
5001
5002
5003
5004
5005
5006
0
200
400
600
Days
b
Observed fraction of stable ASVs
0.8
0.6
0.4
0.2
0.0
*
**
**
*
Overall
5001
5002
5003
5004
5005
5006
Subjects