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. However, it is often difficult to determine which part of the host-associated microbiota is in intimate relationships with the host versus the occasional presence. Mars500 mission was conducted simulating a flight to Mars for 520 days, were 6 crewmembers were kept in isolation and followed standardized diet regimes, constituting a unique experiment to determine, in a longitudinal study design, the composition and importance of the resident stable microbiota vs. a more variable microbiota (changing with time and 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 and evaluating the effect of the three main variables present in the experimental system, the time, the diet and individuality of each subject. Results though showing statistically significant effects of all three variables, highlighted a main

29 contribution of salivary microbiota personalized features, that is an individual-based resilience of
30 the microbiota. Such findings open the way to consider salivary microbiota under the light of a
31 personalized microbiome feature.

32 Introduction

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

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

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

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

70 # Results

71 ## Sample collection

72 Salivary samples were collected by crewmembers during and after the permanence in the
73 isolation facility (Table 1): 42 samples were collected during the first simulated journey from
74 Earth to Mars (seven time-points), 30 samples were collected during the simulated trip back
75 home from Mars to Earth (five time-points), and other 16 samples were collected when
76 crewmembers came back to their normal activities and were followed for additional 200 days
77 (three time-points). Unfortunately, two samples collected in the latter stage gave no good quality
78 DNA and were thus discharged. Crewmembers did not collect salivary samples during the
79 simulated landing on Mars where three of them—which simulated the landing on a separate
80 module—used a different food variant. For this reason the second food variant used was not
81 reported in the work passing from the first food variant (FV) directly to the third food variant
82 (TV). All phases of the mission were reported in Table S1 whereas the full sampling scheme was
83 reported in Figure S1.

84 ## Salivary microbiota composition during the study

85 To inspect how the salivary microbiota reacts in a confined environment, we characterized
86 samples collected during the entire duration of the Mars500 mission (720 days in total) by 16S
87 rRNA gene amplicon sequencing of the variable region V3-V4 (for additional information about
88 the Mars500 mission see Supplementary materials). Amplified sequences formed 1890 amplicon
89 sequence variants (ASVs) with a median number of 172.00 ASVs per sample (ranging from 81 to
90 317). A total of 4,337,540 sequences specifically aligned to an ASV resulting in a sequencing
91 depth ranging from 20,084 to 116,809 and a median value of 47,044 (for additional information
92 about sequence analysis pipeline and the number of sequence obtained in each pre-processing

93 step see Supplementary material, Supplementary Table S2, and Supplementary Figure S2). We
94 estimated accuracy and correlation between technical replicates as described in Supplementary
95 material section. All replicates reported an accuracy higher than 0.96 with a Spearman's rank
96 correlation () that ranged between 0.94 and 0.98 (Supplementary Table S3 and Supplementary
97 Figure S3). Rarefaction curves reached a plateau above 15k reads suggesting an adequate
98 sequencing depth for all samples (Supplementary Figure S4). After accuracy evaluation, one
99 technical replicate for each sample was selected based on the number of reads assigned to ASVs
100 and included in downstream analyses (Supplementary Table S4). Good's coverage estimator
101 ranged between 99.99% and 100.00% across all samples indicating that roughly 0.01% of the
102 reads in a sample came from ASVs that appear only once in that sample (Supplementary Table
103 S5).

104 Roughly, 99% of sequences aligned to variants that came from known bacterial taxa (Table S2).
105 Supplementary Table S6 shows the overall taxonomic composition of samples whereas Figure S5
106 reports the phylogenetic tree reconstructed from ASVs. At phylum level Firmicutes,
107 Bacteroidetes, Actinobacteria, Proteobacteria, and Unknown accounted for more than 94% of the
108 total number of reads assigned to taxonomically annotated ASVs (Figure 1a and Supplementary
109 Table S6). The total bacterial diversity (namely the alpha diversity) remained constant during the
110 mission with no significant differences detected between the isolation period and the follow-up,
111 across different diets, and across subjects (Table S7 and Figure 1 panels b, c, and d). Also time did
112 not impact bacterial diversity as showed by the random mixed model fitted using crewmembers
113 as random intercept which reported a slope value lower than 0.002 (Figure 1e and
114 Supplementary Table S8).

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

116 We inspected differences across samples (namely beta diversity) using non-metric
117 multidimensional scaling (nMDS) on quantitative and qualitative indexes. Samples showed a
118 similar distribution with both index types: Sorensen index and unweighted unifrac distance for
119 qualitative analysis and Bray-Curtis index and weighted unifrac distance for quantitative analysis
120 (Figure 2a). As opposed to alpha diversity, subjects, diets, and time significantly contributed to
121 shape the salivary microbiome with different percentage of variant explained depending on the
122 index but never exceeding 10% of the total variance (Figure 2b). For all diversity indexes (except
123 for the Sorensen index which reported significant effect of subjects) all tested factors reported an
124 homogeneous distribution of dispersion meaning that only the composition of samples varied
125 among groups as highlighted by the permutational analysis of variance reported above.

126 To better explore the effect of time on bacterial diversity we used change-point analysis on both
127 within-subject and between-subjects diversity. Within-subject diversity measured changes in the
128 salivary microbiota of each crewmember through time, whereas between-subjects diversity
129 compared the salivary microbiota of different crewmembers at each time point (Figure 3a and b).
130 Three segments significantly divided within-subject diversity with two change-points at 123 days
131 and 480 days. Between-subjects diversity was not segmented since the overall model gave better
132 results than the segmented one according to the genetic algorithm used during optimization
133 (Figure 3b). The overall between-subjects model had an effect size of -0.00004 which means that
134 after 520 days of experiment the overall diversity decreased by 0.02179. The effect of time on
135 within-subject diversity was indeed higher than the one observed for between-subjects diversity.
136 During the first 123 days the effect modeled was 0.00103 reflecting an average increase of
137 0.12636 for all crewmembers. After the first change-point, within-subject diversity started to

138 decrease with a regression parameter of -0.00064 (average decrease during the second segment
139 of -0.22835). After the second change point, which roughly matched the end of the isolation
140 period (Figure 3c and d), the within-samples diversity started to increase again. At the end of the
141 follow-up period diversity increased again of 0.29474 exceeding the average value detected in
142 the first day of isolation.

143 **Resilience of salivary microbiota**

144 The average abundance of ASVs during the experiment correlates with their persistence—the
145 number of subject in which a given ASV was detected at each time point. Figure 4a shows the
146 increasing trend of log-transformed abundance with an R-squared value of 0.72 (= *10.09*, 95%
147 CI [*9.81*, *10.38*]). Time-resolved clustering produced two groups of ASVs: one—called
148 inconsistent micriobiome (Cluster 1)—included variants detected in a small number of subject at
149 each time point, whereas the other (Cluster 2)—called stable microbiome—included variants
150 detected in the vast majority of subjects during the whole mission (Figure 4a and Figure S7, panel
151 a and b). The inconsistent microbiome showed low average persistence in respect with the stable
152 microbiome but it contained the largest amount of variants (1746 ASVs against 144 of stable
153 microbiome). Unlike stable ASVs, subjects lost and acquired inconsistent ASVs both during and
154 after the isolation period (Figure S7 panel b and c). Stable ASVs were detected in roughly 30% of
155 all subjects at each time point (26 samples on 88) with sporadic losses and acquisitions (Figure
156 4a and Figure S8).

157 We represented the acquisition and loss of bacterial species during the whole mission using
158 networks. At each time point we linked subjects to ASVs detected in their salivary microbiome
159 forming a bipartite network structure which reflected the underlying bacterial community
160 structure. The loss and acquisition of bacterial ASVs was shown in supplementary video S1

161 where green squares represent subjects, red circles represent inconsistent microbiome, and light
162 blue circles represent stable microbiome. As shown in the video, the topology of the networks did
163 not change in time, but at each time point subjects acquire/release bacterial species from/into
164 the environment, except for stable ASVs which are shared by most crewmembers and thus
165 (almost) always present in central part of the network. The number of new edges formed and
166 destroyed passing from one time point to another slightly decreased in time (mixed effect model
167 95% CI for formed edges [-0.08, -0.02] and destroyed edges [-0.08, -0.01] Table S9). The end of
168 the isolation period significantly increased the average number of formed edges (namely
169 acquired ASVs) of 28 but the trend was still negative (Figure 4b). The number of formed edges
170 was independent from the number of lost edges (Spearman's = -0.04; p-value = 0.750). The
171 salivary microbiome structure did not change both during and after the isolation period
172 reporting a similar network topology even if we exclude lost ASVs (Figure 4c).

173 **Discussion**

174 Mars500 experiment evaluated the temporal dynamics of the human salivary microbiota in a
175 controlled and confined environment. All samples from the 6 (male) crewmembers included
176 Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, and Unknown as main phyla and
177 conserved their taxonomic composition along time and among individuals, as previously
178 reported by other studies (see for instance[7,9,18,19]). Despite this conservation, external factors
179 significantly influenced the salivary microbiome composition but their influence was restricted
180 to a low percentage of the community (less than 10% of total variance explained). We found that
181 a great number of species intermittently passed through the salivary microbiome but they
182 neither affected the overall microbiome structure nor its taxonomic composition. Despite their
183 number, transient species struggle to thrive in human saliva reporting a low abundance during

184 the whole experiment. On the other hand, a hundred bacterial taxa dominated the salivary
185 microbiome composition of all crewmembers rarely changing host and abundance even when
186 isolation period ended. The isolation time affected bacterial diversity of single individuals but it
187 did not alter the microbiome of the whole crew. The bacterial diversity of crewmembers
188 decreased—if compared within consecutive time points of the same individual—but the effect
189 ended immediately after the isolation period when it started to increase again. External
190 perturbations, impossible to control outside the isolation facility, modulated the salivary
191 microbiome composition when crewmemebrs got out the isolation facility and started eating
192 different types of food, getting in touch with other people, or simply visiting different places.

193 This finding suggests that sharing the same confined environment—and possibly following the
194 same diet regime with few variation—imbalanced the ecology of the salivary microbiota while
195 reducing its complexity. Unfortunately, we do not know if this effect could—positively or
196 negatively—affect the health of the hosts but, according to the holobiont theory of evolution, a
197 loss of strains in the microbiome correspond to a loss of putative members that could be selected
198 if environmental conditions changes.[3,4,20] Extending this concept we could affirm that a
199 depauperate microbiome is less reactive to sudden changes in external conditions weaken host
200 defenses. However, between-subjects differences remained unaltered suggesting that individuals
201 follow somewhat independent dynamics of their salivary microbiota (i.e. personalized dynamic).
202 The same evidence indicates that crewmembers, though sharing the same environment, did not
203 exchange their salivary microbiota, leading to hypothesize that quite stable personal salivary
204 microbiota features are present in humans and confirming the between-subject effect. The fecal
205 microbiota of the same crewmembers showed an increasing trend of similarity among subjects,
206 especially in relation to a sharing of rare taxa,[17] indicating that salivary microbiota has a more

207 pronounced personalization than fecal microbiota. Indeed, from our data there are no evidences
208 of an ecological successions, as those shown in the fecal microbiota of the same crewmembers.

209 Even if factors such as diet and time influenced the salivary microbiome composition of
210 crewmembers, most details of the Mars500 missions are unknown. The mission was a military
211 experiment and several outcomes are still sealed. The composition of diets for example is
212 unknown and thus any speculation on the effect of particular food intake would not be grounded.
213 Since participants to the study must remain isolated from the outer world, they collected samples
214 inside the isolation facility affecting the reproducibility of the experiment and influencing the
215 microbiota variation due to different subjects. Despite these limitations the use of a high-
216 resolution technique, such as 16S rRNA gene sequencing, allowed us to detect key features of
217 human salivary microbiome under a condition that is almost impossible to replicate. The
218 complete isolation of the participants of the mission made possible the first observation of
219 salivary microbiome composition minimizing the effect of external perturbations.

220 In conclusion, the reported longitudinal analysis of human salivary microbiota confirmed the
221 stability of the microbiota over time and suggested the presence of resilient personalized
222 taxonomic features, which may deserve further attention in the future. This study allowed clearly
223 to determine the contribution of a stable and confined environment, as that of Mars500
224 experiment, in reducing the microbiota diversity and to show the effect of a controlled diet on
225 salivary microbiota.

226 # Methods

227 ## The Mars500 experiment

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

242 ## Collecting salivary samples

243 Saliva samples were collected individually, based on the scientific protocol and pre-confinement
244 training, with 5ml sterile vials (Nalgene V5257-250EA), Upon completion of the saliva sampling
245 all samples from one sampling event were put into the hatch. After that, they were removed by
246 the responsible person of the IBMP and stored at -80°C. After being stored at -80°C in the
247 laboratories of the IBMP for periods of at least 4 days up to 6 months, the samples were sent via
248 World Courier. Shipping from Moscow to the University of Tuscia, Viterbo – Italy. The shipment

249 was performed in three batches on dry ice to avoid repeated freeze-thaw cycles which lead to
250 reduction of microbial viability. Upon arrival, samples were kept at -80°C until processing.

251 **Sequencing of salivary samples**

252 DNA was extracted from salivary samples stored at -80°C using a conventional bead-beating
253 protocol (DNeasy PowerSoil Kit, Mobio). After fluorimetry quantification (Qubit), 20 ng of
254 environmental DNA were used as template for amplification of 16S rRNA gene using V3-V4
255 primers (341F and 785R) as previously reported.[21,22] Libraries were constructed and sequenced
256 on a MiSeq apparatus[23] (Illumina) by BMR Genomics (Padua, Italy). Sequences are deposited at
257 ENA database under the accession ERP119217.

258 **Amplicon sequence variant reconstruction**

259 The DADA2 pipeline (version 1.14.1) was used to reconstruct amplicon sequence variants (ASVs)
260 from illumina reads 24. Both ASV reconstruction and statistical analyses were performed in the R
261 environment version 3.4.3 (http://www.R-project.org). For a complete description of all the step
262 performed see Supplementary material section. Briefly, primers used for V3-V4 amplification
263 were detected and removed using cutadapt version 1.15.[25] Low quality reads were discarded
264 using the `filterAndTrim` function with an expected error threshold of two for both forward and
265 reverse read pairs. Denoising was performed using the `dada` function after error rate modelling
266 (`learnErrors` function). Denoised reads were then merged discarding those with any
267 mismatches and/or an overlap length shorter than 20bp ('mergePairs' function). Chimeric
268 sequences were removed using the `removeBimeraDenovo` function. Taxonomical classification
269 was performed using DECIPHER package version 2.14.0 against the latest version of the pre-
270 formatted Silva small-subunit reference database (SSU version 132 available at:
271 http://www2.decipher.codes/Downloads.html).[26,27] All sequences classified as chloroplasts,

272 mitochondria, Archaea and Eukarya were removed. A summary of retained reads in each step is
273 reported in Table S1 and in Figure S1.

274 Diversity estimation

275 Bacterial diversity in each sample was computed using inverse Simpson index as implemented in
276 the `diversity` function of vegan package. Differences according to crewmembers, permanence in
277 the isolation facility, and food variants were inspected using one-way analysis of variance
278 (ANOVA). The effect of time was modeled using linear mixed models with fixed slope and random
279 intercept. Since alpha diversity was measured multiple times on the same statistical units,
280 crewmembers were used as random intercept factor.

281 Diversity across samples was inspected using different approaches. Qualitative and quantitative
282 indexes were used to infer pairwise distances between samples. Qualitative indexes are binary
283 indexes which take into account presence/absence of species to compute distances between
284 samples whereas quantitative indexes are mainly based on the abundance of species.[28] Sorensen
285 index[29] and un-weighted UniFrac distance[30] were used as qualitative indexes whereas Bray-
286 Curtis dissimilarity[31] and weighted UniFrac distance[30] were used as quantitative indexes. UniFrac
287 distances were computed using the `distance` function of the phyloseq R package version 1.30.0[32]
288 whereas Sorensen and Bray-Curtis dissimilarity indexes were computed using the `vegdist`
289 function of the R package vegan verison 2.5-6.[33] Differences between salivary microbiome
290 composition of the same crewmember at consecutive time-points (within-subject diversity) were
291 computed using the `TBI` function of the adespatial R package verison 0.3-8.[34] Packages `vegan` and
292 `adespatial` use the same definition of Sorensen distance, defined as:

$$\frac{(+\ -2)}{(+)}$$

293 where A and B are the numbers of ASVs on compared samples, and J is the number of the ASVs
294 shared by both samples. Qualitative indexes rely on the assumption that all taxa are equally
295 contributing to bacterial diversity independently from their abundance. For this reason, even
296 extremely rare taxa may be relevant in shaping sample distribution. To relax this assumption
297 ASVs detected in less than 5% of samples (4 samples overall) with an abundance lower than 10
298 were filtered out before diversity calculation.

299 Distances across samples were reported using non-metric multidimensional scaling (nMDS) as
300 implemented in the `metaMDS` function of the vegan package, with 300 random starts and
301 monotone regression.[35,36] To test the effect of food variants, crewmembers, and time in shaping
302 the salivary microbiome we used permutational multivariate analysis of variance on distance
303 matrices obtained above (`adonis2` function of the vegan package with 1,000 permutations). The
304 proportion of sum of squares from the total (namely the R^2 value of permutational analysis)
305 was used to report the percentage of variance explained by each factor included in the analysis.
306 Before testing for differences in bacterial composition among groups is advisable to make sure
307 that groups are homogeneously dispersed, otherwise permutational tests (such as `adonis`) may
308 report significant results entirely due to uneven dispersion. To distinguish between actual
309 differences in composition or differences due to dispersion we used the `betadisper` and `anova`
310 functions (`vegan` package). P-values obtained were corrected using Benjamini & Hochberg
311 correction (also known as false discovery rate)[37]. To avoid possible biases induced by uneven
312 sequencing depths, read counts were scaled using DESeq2 before diversity calculation (`counts`
313 function) 38. Scaled counts were additionally transformed using the square root of the Wisconsin
314 double standardized counts (`wisconsin` function of the vegan package).

Influence of time on bacterial diversity

Salivary microbiome may be affected by several factors. Sharing the same environment for a prolonged period of time may alter the composition of salivary microbiome at different levels. The bacterial composition may be altered within the same individual taken at consecutive time points but even between multiple individuals at each time point. To inspect both of these components, bacterial diversity within and between subjects, calculated as reported above, was modeled through time. Change-point analysis was used to identify specific time points which led to a decrease/increase of diversity. The optimal positioning and number of change-points for each crewmember was identified using a non-parametric cost function as implemented in the `cpt.np` function of the changepoint.np R package, version 1.0.1.[39] The pruned exact linear time algorithm (PELT)[40] was used to detect temporal changes in diversity within the same subject and between different subjects. The PELT algorithm searches for an optimal solution by minimizing the cost of different segmentation. We used the modified Bayes information criterion penalty term (MBIC)[41] as penalty function for cost minimization by the algorithm. Since PELT algorithm is exact, a solution is always found for each time series so, to avoid inflation of change-points due to the presence of data coming from six different subjects, a genetic algorithm was used to fine-tune the analysis. All change-points detected were used as starting point of the genetic algorithm and a fitness function was defined as:

$$\frac{1}{\sum_{=1} ()}$$

where, () stands form the root mean square error the generalized linear model constructed on the segment and is the number of segments defined by change-point analysis. High error corresponds to a low fitness value whereas low error corresponds to a high fitness value. At each

336 step of iteration, the genetic algorithm will keep segmentation that lead to linear models with a
337 low RMSE and discard those leading to high error models. We implemented this algorithm using
338 the R package GA[42] verison 3.2 with a population size of 200. Generalized linear model were
339 fitted using crewmembers as random intercept and p-value were computed with the
340 Satterthwaite's degrees of freedom method as implemented in the package lmerTest[43,44] version
341 3.1-1.

342 We assessed the persistence of the salivary microbiome across subjects using dynamic time
343 warping algorithm implemented in the dtwclust R package (version 5.5.6)[45]. At each time point
344 the number of subjects in which a given ASV was detected (namely the ASV's persistence) was
345 reported together with its abundance. Persistence matrix was scaled and centered before
346 clustering. Centering was performed by subtracting the mean of each ASV from their persistence
347 whereas scaling was performed by dividing each value by its standard deviation. The relation
348 between persistence and abundance was tested by fitting a linear model ('lm' function of R stat
349 package). All detail s about clustering and modelling were reported in Supplementary methods.
350 Persistence across subjects was also used for network construction: at each time point we
351 constructed a bipartite network by linking subjects with ASVs that were present in their salivary
352 microbiota. Subjects were represented using squared nodes whereas ASVs were represented
353 using round circles colored according to the groups defined above. Doing so, we generated twelve
354 bipartite, acyclic, and undirected networks representing the salivary microbiota of all subjects at
355 different time points. The network R package (version 1.16.0) was used for network
356 reconstruction whereas the package ggnetwork (version 0.5.8) was used for plotting.[46,47] The
357 effect of time and the end of the isolation period on the number of formed/destroyed edges were
358 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.

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 microbiome.

Days		SE			
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. , regression parameter (slope of the model); SE, standard error; , t-value (also known as “standardized” regression parameter); , degrees of freedom; , p-value.

Figures

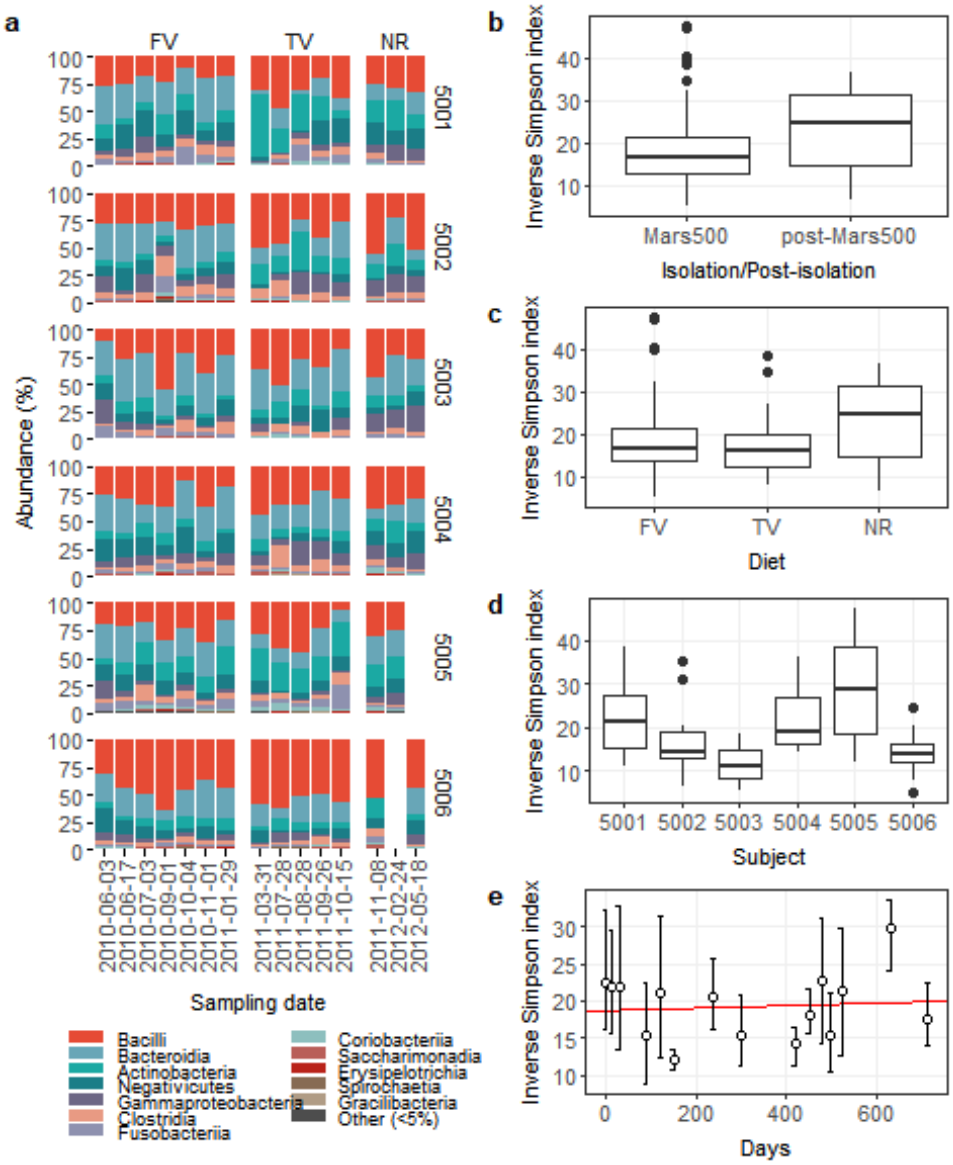

Figure 1: Salivary microbiota diversity along Mars500 mission and follow-up. a) Distribution of the main bacterial classes. Panels were divided according to crewmembers (vertically) and diets (horizontally). ASVs with a relative cumulative frequency lower than 5% in all samples were collapsed into a 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.

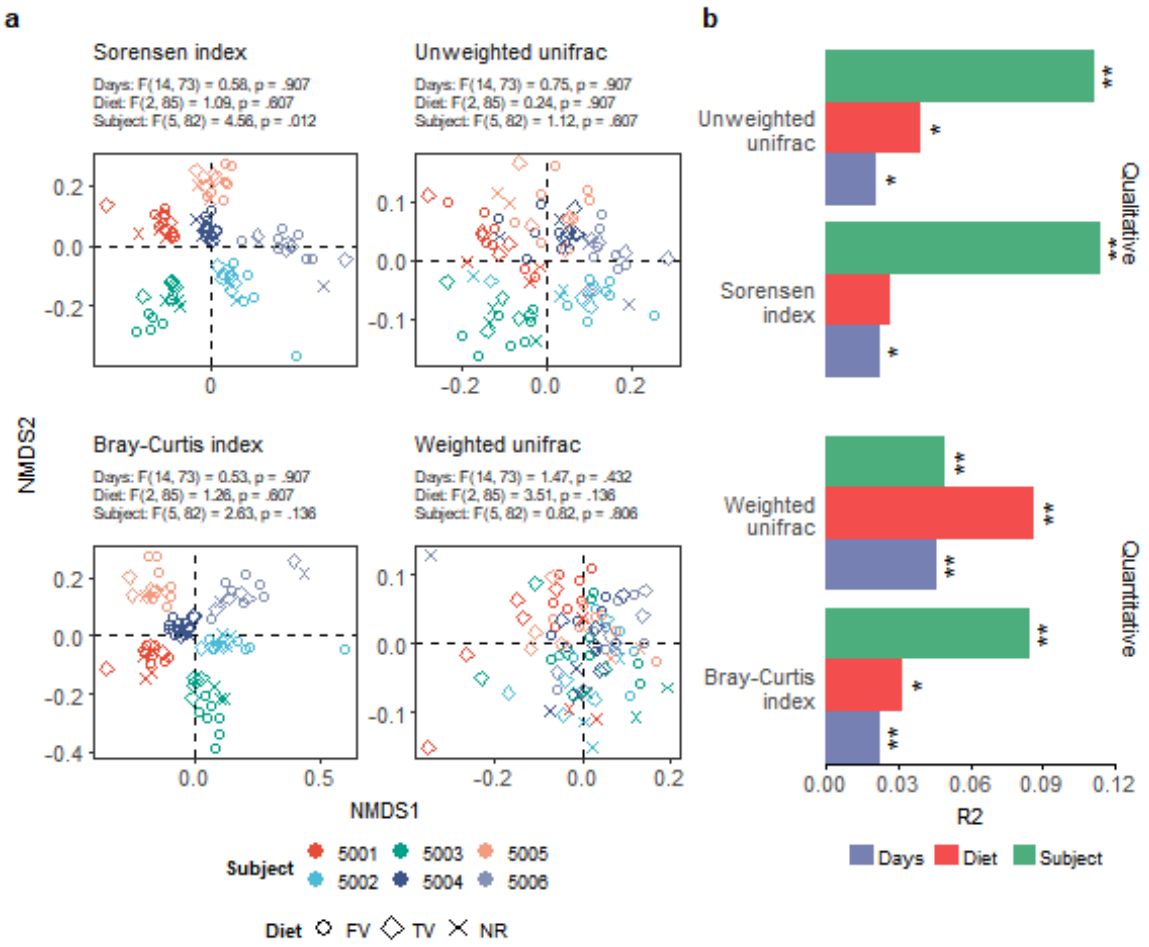

Figure 2: Microbial assemblage variation according to diet, crewmembers, and time. a) Non-metric multidimensional scaling based on different beta-diversity indexes (reported on the top of each panel). Samples were colored according to crewmembers whereas the point shape represents the type of diet (FV, 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.

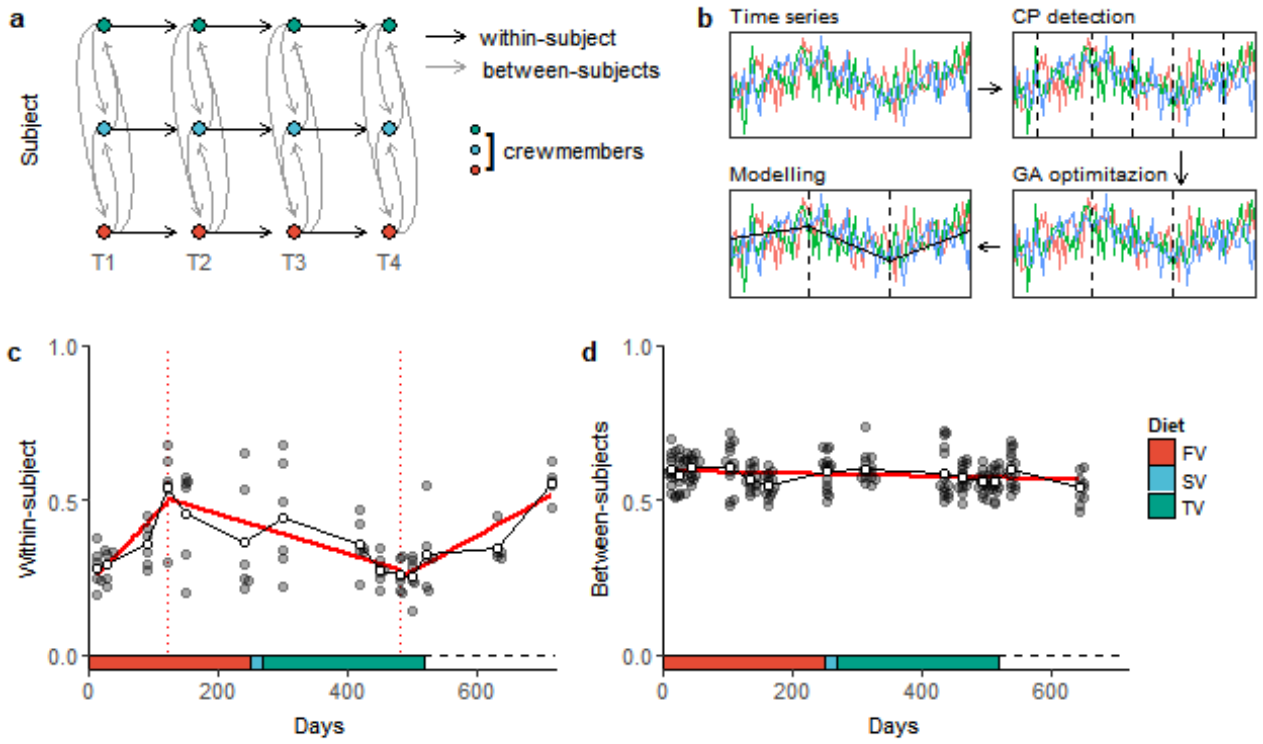

Figure 3: Crewmembers' salivary microbiome composition in time. a) Sorensen index has been used to inspect distances between and within-subject during isolation and follow-up. Between-subjects diversity was computed by comparing the salivary microbiome of each subjects at each timepoint (gray arrows); within-subject diversity was computed by comparing the salivary microbiome of the same subject over time (black arrows). b) Change-point analysis revealed changes in salivary microbiome 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.

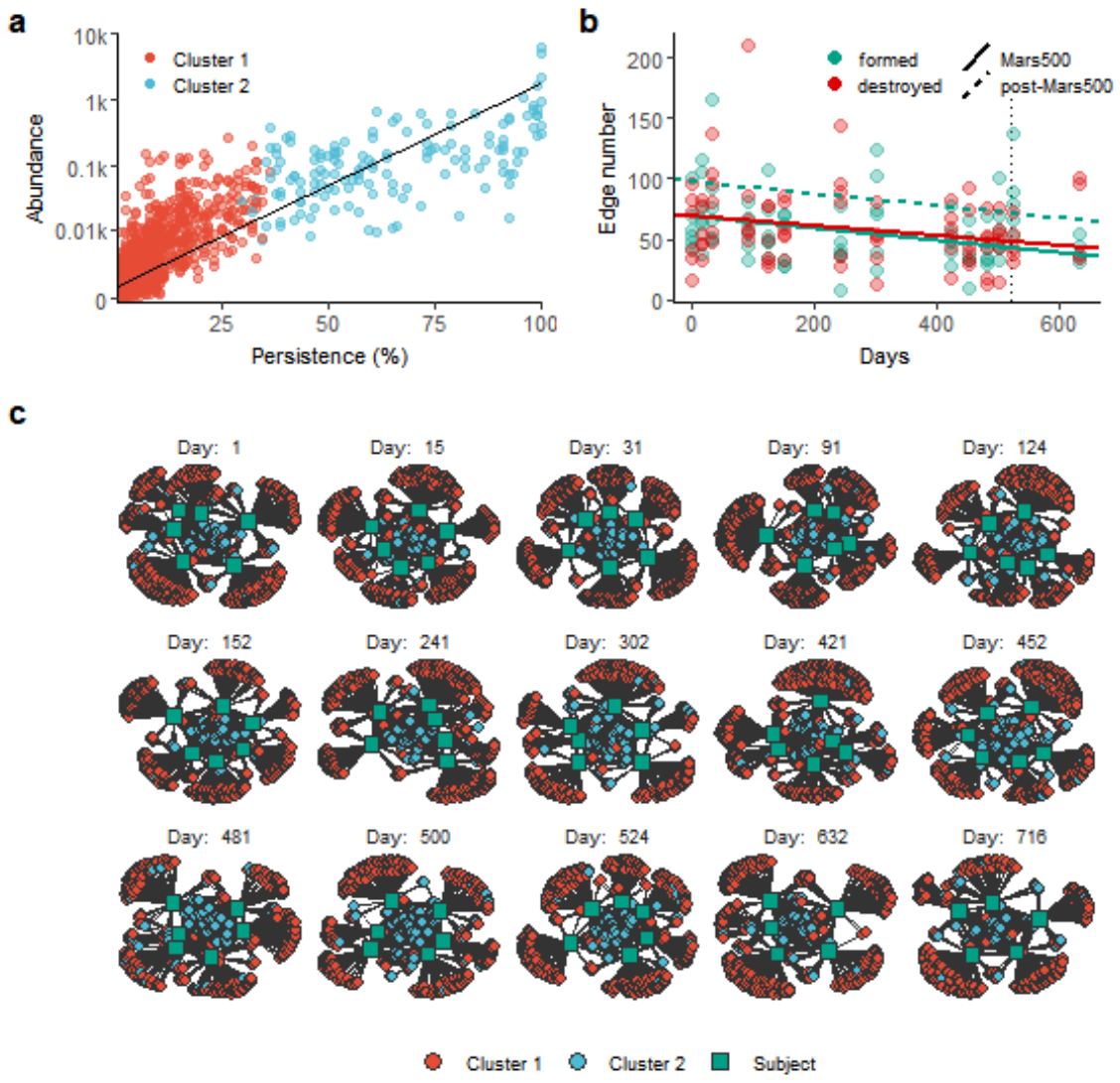

Figure 4: Bacterial community structure in time. a) Persistence and abundance of ASVs detected. Persistence was expressed as the number of subject in which an ASV was detected whereas abundance was expressed as log-normalized number of reads assigned to that ASV (the black line represent results of the 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.

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.

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

19. 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).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

37. 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).

38. 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).

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

40. 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).

41. 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).

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

43. 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).

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

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

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

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