parafac4microbiome: Exploratory analysis of longitudinal microbiome data using Parallel Factor Analysis

G.R. van der Ploeg[1], J.A. Westerhuis[1*], A. Heintz-Buschart[1], A.K. Smilde[1]

[1] *Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, The Netherlands*

*Author for correspondence: Johan Westerhuis j.a.westerhuis@uva.nl

Geert Roelof van der Ploeg: https://orcid.org/0009-0007-5204-3386

Johan Westerhuis: https://orcid.org/0000-0002-6747-9779

Anna Heintz-Buschart: https://orcid.org/0000-0002-9780-1933

Age Smilde: https://orcid.org/0000-0002-3052-4644

Abstract

Background: Recently, studies that investigate microbial temporal dynamics have become more frequent. In a longitudinal microbiome study design, microbial abundance data are collected across multiple time points from the same subjects. In this context, exploratory analysis of longitudinal microbiome data using Principal Component Analysis is insufficient because the study design is not fully utilized. Indeed, the study design can be used to create a three-way data array where the rows constitute the subjects, the columns contain the microbial abundances, and the third dimension contains the time points. Organising the data this way enables multi-way methodologies that take full advantage of the study design. In this work we present Parallel Factor Analysis (PARAFAC) as a method to explore longitudinal microbiome data using three exemplary studies.

Results: In the first example study, a long time series of experimental in vitro microbiomes, we show that PARAFAC can identify the main time-resolved variation in longitudinal microbiome data. In the second example, a longitudinal infant gut microbiome study, we show that PARAFAC can find differences between subject groups and enhance comparative analysis despite a moderate amount of missing data. In the third example, a longitudinal gingivitis intervention study of the oral

microbiome, we show that PARAFAC helps identify microbial groups of interest using a post-hoc clustering approach.

Conclusions: We show that Parallel Factor Analysis is an applicable method for longitudinal microbiome data analysis across a wide range of microbial environments. The analyses and the example datasets with the resulting figures are implemented in the R package `parafac4microbiome`, which is available online at https://github.com/GRvanderPloeg/parafac4microbiome.

Key Words

Multi-way, PARAFAC, Tensor Factorization, Decomposition, Dynamics

Background

The role of the microbiome in biological systems and in disease has been thoroughly explored [1–3]. However, most studies are cross-sectional while microbial temporal dynamics are required to understand host-microbiome and microbe-microbe interactions [4, 5]. Recently, studies that investigate these dynamics have become more frequent [6–8].

In a longitudinal microbiome study design, microbial abundance data are collected across multiple time points from the same subjects [9–11]. These data are often stored in a two-way data table, where the rows constitute the measurements (samples from multiple individuals over time), and the columns contain the microbial abundances (**Figure 1A**). Two-way ordination methods like Principal Component Analysis [12, 13] are then employed to explore the data (**Figure 1B**). However, this often results in hard-to-interpret visualizations because these methods treat every measurement as a separate object to be ordinated (**Figure 1D, E**). As such, there is a need for exploratory methods that take the crossed longitudinal study design into account.

The longitudinal study design can be used to create a three-way data array where the rows constitute the subjects, the columns contain the microbial abundances, and the third dimension contains the time points (**Figure 1A**). Three-way ordination methods like Parallel Factor Analysis (PARAFAC; [14, 15]) can then be used to explore the data (**Figure 1C**). By making full use of the study design, the model can identify subject groups as well as their time profiles (**Figure 1D, F**).

In this work we present Parallel Factor Analysis as a method to explore longitudinal microbiome data. We first give a mathematical description of the PARAFAC method and its assumptions,

58 followed by a brief review of longitudinal microbiome data processing needed to meet these
59 assumptions and the principles of PARAFAC model selection. We then highlight the benefits of
60 making a PARAFAC model using three exemplary datasets. In the first example study [6] we
61 highlight the time-resolved variation that PARAFAC models. In the second study [8], we highlight
62 how PARAFAC can find differences between subject groups and enhance comparative analysis
63 despite a moderate amount of missing data. In the third study [7] we highlight a post-hoc clustering
64 approach to identify microbes of interest with PARAFAC. The analyses and the example datasets
65 with the resulting figures are implemented in the R package `parafac4microbiome`, which is
66 available online at https://github.com/GRvanderPloeg/parafac4microbiome.

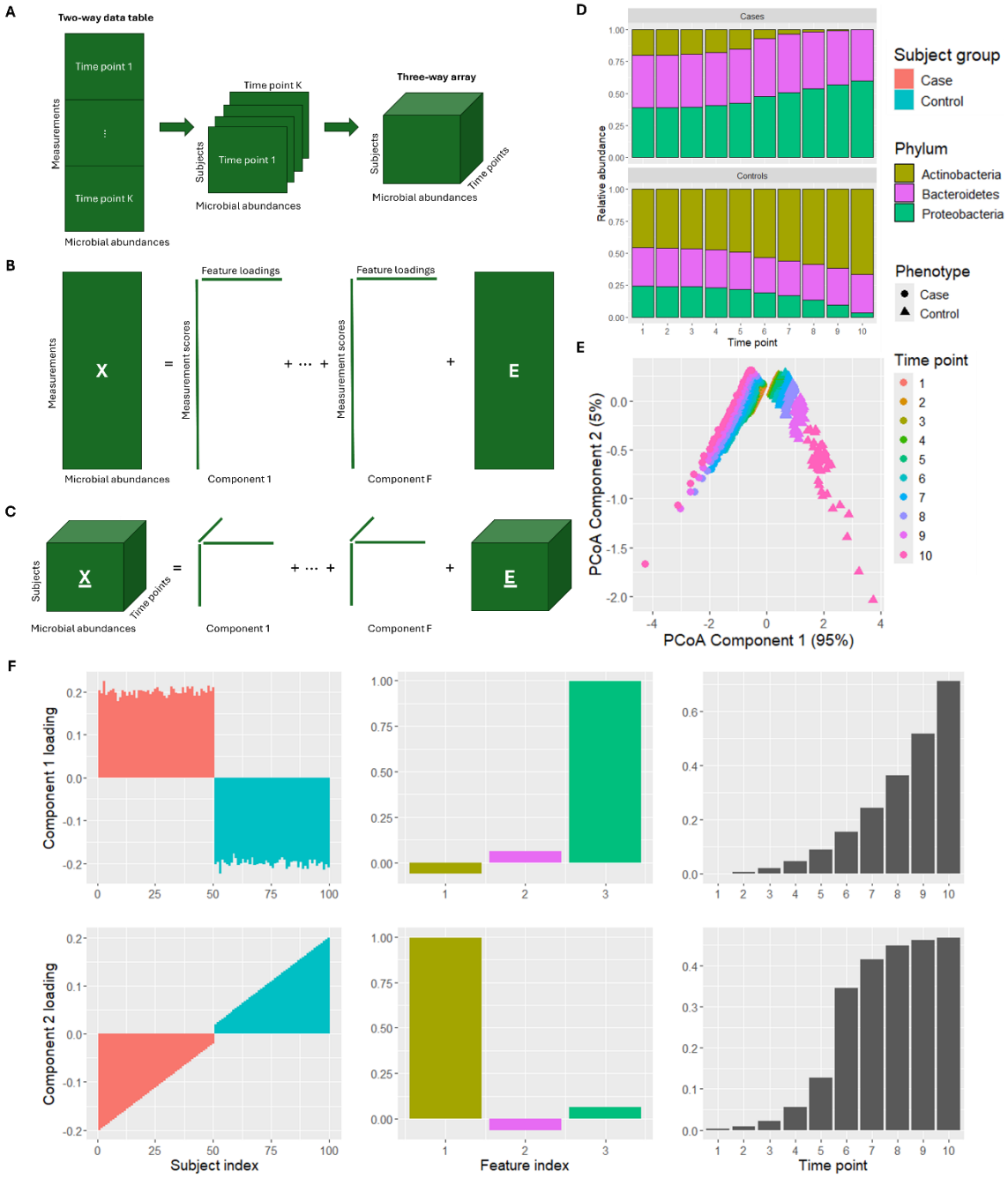


Figure 1: (A) Longitudinal microbiome data is often presented as a two-way data table with measurements (i.e. one sample for every subject-time point combination) in the rows and microbial abundances in the columns. This kind of data can be converted to a three-way array by putting all measurements for each time point in the third dimension. (B) Principal Component Analysis decomposes a two-way data table into components with scores for the measurements and loadings for the microbial relative abundances. This does not take the longitudinal study design into account. (C) Parallel factor analysis decomposes a three-way data array into components with three loading vectors each: one for the subjects, one for the microbial relative abundances, and one for time. (D) In a mock intervention dataset, we have obtained count data of three phyla (Actinobacteria, Bacteroidetes and Proteobacteria) corresponding to 50 case and 50 control subjects across 10 time points. At time point 1 all subjects look very similar to each other. At time point 10 the case subjects are depleted of Actinobacteria and enriched in Proteobacteria, while for the control subjects the opposite is true. The depletion and enrichment of the Actinobacteria at time point 10 varies substantially within the case and control groups, respectively. (E) Principal Coordinate Analysis of the mock data with Aitchison distance. While the subject groups and time points are separated, the time profiles of each group are unclear. (F) Using Parallel Factor Analysis, we can identify both the subject groups and the time profiles present in the data. In component 1 the exponential growth of Proteobacteria in the case subjects is identified. In component 2 the sigmoid shaped growth curve of Actinobacteria in the control subjects is identified. The model describes time-resolved differences in microbial abundances between the subject groups.

Methods

Notation and definitions

We briefly define the mathematical notation that will be used throughout this paper. Scalars are indicated by lower-case italics such as a. Vectors are indicated by bold lower-case characters such as $\boldsymbol{b}$. Two-way matrices are indicated with bold capitalized characters such as $\boldsymbol{X}$. Underlined bold capitalized characters are used for three-way arrays such as $\underline{\boldsymbol{X}}$. The letters I, J, and K are reserved to indicate the dimension of the subject, microbial abundance, and time mode, respectively. Hence the element for subject $i \in (1 \dots I)$, microbial abundance $j \in (1 \dots J)$ at timepoint $k \in (1 \dots K)$ of a three-way array $\underline{\boldsymbol{X}}$ is called x_{ijk}. We create the data cubes of microbial relative abundances such that the subjects are in the first mode, the microbial taxa are in the second mode, and time points are in the third mode. We will not distinguish between the terms 'factor' and 'component', nor between the terms 'way' and 'mode'. Generally, we will use the words 'component' and 'mode' throughout this paper.

Mathematical definition of Parallel Factor Analysis

Parallel Factor Analysis (PARAFAC) [14, 15] is a factor decomposition method that generalizes Principal Component Analysis (PCA) to multi-way. First, we give a brief mathematical summary of PCA to make the generalizing principle of PARAFAC clearer. We then describe the algorithmic implementation of this method and its properties.

A longitudinal microbiome dataset is often represented as a two-way matrix $\boldsymbol{X}$ $(IK \times J)$ with all time points for all subjects as rows and the microbial abundances as columns (**Figure 1B**). For convenience, we will refer to a subject-timepoint combination as a measurement $m \in (1 \dots M)$ where $M = IK$. A PCA model of this two-way dataset $\boldsymbol{X}$ $(M \times J)$ is then given by the two matrices $\boldsymbol{T}$ $(M \times F)$ and $\boldsymbol{P}$ $(J \times F)$ containing the scores and loadings of all F components respectively. The scores describe the differences between the measurements, and the loadings describe the microbial abundances that are different. A PCA model of such a two-way matrix can be formulated elementwise as

$$x_{mj} = \sum_{f=1}^{F} t_{mf} p_{jf} + e_{mj} \tag{1}$$

with x_{mj} the matrix element corresponding to measurement m and feature j, t_{mf} the score of measurement m in component f, p_{jf} the loading of feature j in component f and e_{mj} the residual of the matrix element. PCA is a bilinear model as we are using F sets of two vectors to approximate the dataset $\boldsymbol{X}$. This also reveals the main issue at hand: the measurements M are ordinated without taking the connection to the subjects I and time points K into account.

A three-way data array $\underline{\boldsymbol{X}}$ is defined as having I subjects in the rows, J microbial abundances in the columns and K time points in the third dimension (**Figure 1C**). A PARAFAC model of a three-way data array $\underline{\boldsymbol{X}}$ $(I \times J \times K)$ is given by three matrices $\boldsymbol{A}$ $(I \times F)$, $\boldsymbol{B}$ $(J \times F)$ and $\boldsymbol{C}$ $(K \times F)$ containing the loadings of all F components corresponding to the subject, feature, and time mode, respectively. The element-wise definition of the model is formulated as

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk} \tag{2}$$

with x_{ijk} the matrix element corresponding to subject i, feature j at time point k, a_{if} the loading of subject i of component f, b_{jf} the loading of feature j of component f and c_{kf} the loading of time point k of component f and e_{ijk} the residual of the matrix element. In the elementwise definition, the generalization of PCA is readily visible: the loading for the time mode c_{kf} has been added to the calculation of element x_{ijk} to represent the time factor. PARAFAC is a trilinear model as we are using F sets of three vectors to approximate the data array $\boldsymbol{X}$.

Algorithmic implementation of Parallel Factor Analysis

The objective of the algorithmic implementation of PARAFAC is to minimize the sum-of-squares of the residuals through the loss function

$$L(a_{11}, a_{12}, \dots, c_{KF}) = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(x_{ijk} - \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} \right)^2 \tag{3}$$

Various implementations of estimating the PARAFAC loading matrices exist [16]. In this paper we estimate the PARAFAC loading matrices through an iterative alternating least-squares algorithm (ALS; [15, 17]), as this implementation is well-documented and robust towards missing values or time points (**Table 1**).

Table 1: Algorithmic implementation of the alternating least squares based PARAFAC model estimation.

Parallel Factor Analysis (PARAFAC) ALS algorithm for a given three-way array $\underline{\boldsymbol{X}}$
(0) Decide on the number of components F
(1) Initialize $\boldsymbol{B}$ and $\boldsymbol{C}$
(2) Estimate $\boldsymbol{A}$ from $\underline{\boldsymbol{X}}, \boldsymbol{B}$ and $\boldsymbol{C}$ such that $L(a_{11}, a_{12}, \dots, c_{KF})$ is minimized
(3) Estimate $\boldsymbol{B}$ from $\underline{\boldsymbol{X}}, \boldsymbol{A}$ and $\boldsymbol{C}$ such that $L(a_{11}, a_{12}, \dots, c_{KF})$ is minimized
(4) Estimate $\boldsymbol{C}$ from $\underline{\boldsymbol{X}}, \boldsymbol{A}$ and $\boldsymbol{B}$ such that $L(a_{11}, a_{12}, \dots, c_{KF})$ is minimized
(5) Continue to step (2) until convergence.

Various approaches for initialization of $\boldsymbol{B}$ and $\boldsymbol{C}$ have been proposed and are documented elsewhere [18]. In this paper, we used randomly initialized vectors to determine the appropriate number of components and singular value decomposition-based initialized vectors to investigate the stability of the loadings.

We use the PARAFAC algorithm from the multiway R package (version 1.0-6; [19]). With our default settings, the model is converged in step (5) when the change in sum-of-squares is lower than 1e-6 or after 500 iterations of steps 2-4. Fitting a model for any of the processed datasets in this paper takes less than a minute of computational time on an average computer (Microsoft Windows 11 Home v10.0.22631 with an 12th Gen Intel® Core™ i5-12400F running 2500 MHz with 6 cores or 12 logical processors and 16 Gb RAM).

PARAFAC model constraints and properties

A PARAFAC decomposition is essentially unique, meaning that the solution is unique apart from trivial rescaling and permutations of the factor scores, which is a distinct advantage over Principal Component Analysis [15, 20, 21]. Additionally, in PCA the components are orthogonal while in PARAFAC the components are non-orthogonal by default. The orthogonality constraint in PCA results in models that force interpretation in terms of orthogonal factors which do not always align with the underlying biology, whereas the absence of such a constraint for PARAFAC results in models that can describe processes that are not necessarily independent. Non-orthogonality is an advantage because it is more likely a realistic representation of the biological system of interest. The downside of non-orthogonality is that components can be very similar to each other. As such, interpretation should be done with care: examining one component at a time can lead to wrong conclusions and plotting the loadings of two components against each other requires a

transformation (**Supplementary Methods**). Additional constraints such as orthogonality and non-negativity of the loadings can be imposed by the user and are listed elsewhere [18, 19]. In this paper we will not constrain the PARAFAC models.

Processing of longitudinal microbiome data

In this paper we perform the following four processing steps for each dataset to meet the assumptions of the PARAFAC modelling procedure: centred log-ratio transformation, feature filtering, centring and scaling. In this section we outline these operations for multi-way data arrays. In the supplied R package, the function processDataCube() performs each of these steps according to user specifications.

Centred log-ratio transformation (CLR) is a popular method to create a version of the data that can be ordinated and interpreted through Euclidean distances [22]. This transformation approach deals with the compositionality inherent to microbiome data and bring the data closer to a normal distribution [23]. However, it is hampered by the zero-inflation commonly observed in microbiome data because zeroes are not defined in log space [24]. Here we choose to CLR transform the datasets after adding a pseudo-count of 1 to all counts.

Due to the high dimensionality of the microbiome data, filtering of the features is a requirement to model the data appropriately and quickly [25, 26]. However, in longitudinal microbiome data the complete time trajectory of a feature should be considered prior to filtering. As such, we define the sparsity of a feature as the fraction of measurements where the feature has zero counts. A sparsity threshold can then be established to select features that are common, thus controlling the number of zeroes in the data.

Centring is used to remove offsets in the data [27, 28]. This can be done in three different directions for three-way arrays: across the subject mode, across the feature mode and across the time mode. Centring across the subject mode removes the averages from every combination of microbial abundance and time point. Centring across the feature mode removes the averages from every combination of subject and time point. Centring across the time mode removes the averages from every combination of subject and microbial abundance. Both centring across the subject mode and across the feature mode affects the interpretation of the time mode as deviations from the average profile will be modelled. Since we expect individual offsets to be present in the exemplary studies,

191 we centre across the subject mode. As a result, the PARAFAC models will describe deviations from
192 the average time profile.

193 Scaling is used to change the importance attached to different parts of the data [27, 28]. For three-
194 way arrays, this involves scaling a two-way matrix to standard deviation one. As such, there are two
195 possibilities for three-way arrays: scaling within the subject mode and scaling within the feature
196 mode. Scaling within the subject mode makes the standard deviation for all counts across all subjects
197 and microbial abundances per time point equal to one. This makes all time points equally important
198 for the modelling procedure. Scaling within the feature mode makes the standard deviation across
199 all subjects and time points per microbial abundance equal to one. This makes all microbial
200 abundances equally important for the modelling procedure. Since we are interested in groups of
201 microbiota with different time profiles, we scale within the feature mode.

202 *Example dataset: Fujita et al., 2023*

203 The first example is a long time series of experimental in vitro microbiomes [6]. Culture conditions,
204 microbiome sampling, 16S rRNA gene amplicon sequencing and data pre-processing are described
205 in the original publication. We analyse only one experimental set-up, the aquatic microbiome
206 culture in a peptone medium, which was performed in eight replicates and measured daily for 110
207 days.

208 The amplicon sequence variant (ASV) data and taxonomic information were processed using the
209 `parafac4microbiome` package (v1.0.0) in R (v4.3.2; [29]). The supplied dataset is a three-way array
210 of counts of size 8 replicates x 28 microbial (ASV) abundances x 110 time points. Features were kept
211 if they had ≤99% sparsity across the dataset (**Supplementary Figure 1**). This step resulted in 23
212 features being selected. We performed a centred log-ratio transformation with a pseudo-count of 1
213 to correct for compositionality [22, 23]. The data was then centred across the subject (here: replicate)
214 mode and scaled within the feature mode.

215 *Example dataset: Shao et al., 2019*

216 The second example is a longitudinal faecal microbiome study with samples collected from 314
217 vaginally born and 282 caesarean-section born infants [8]. Sampling was performed at four time
218 points after birth: day 4, 7, 21, and a follow-up sample was taken between 4 to 12 months (referred
219 to as the infancy period). Recruitment, microbiome sampling and 16S rRNA gene amplicon
220 sequencing and pre-processing are described in the original publication.

Pre-processed sequencing data were obtained from ENA (accession numbers ERP115334 and ERP024601). Taxonomic classification was performed using mOTUs (v2.5.1; [30]). The marker-gene based operational taxonomic unit (mOTU) data and taxonomic information were processed using the `parafac4microbiome` package (v1.0.0) in R (v4.3.2). The supplied dataset is a three-way array of counts of size 395 subjects x 959 microbial (mOTU) abundances x 4 time points, keeping missing measurements as a row of NAs. This allows PARAFAC to impute the missing values. In total, 62.5% of the data was missing (**Supplementary Tables 1-2**). mOTUs were kept if the sparsity in either birth mode group was ≤90% (**Supplementary Figure 5**). This conserves biologically relevant mOTUs that only appear in one group and resulted in 91 of the 959 mOTUs being selected. The data was then centred log-ratio transformed using a pseudo-count of 1 to correct for compositionality [22, 23]. Subsequently, the data was centred across the subject mode and scaled within the feature mode.

Example dataset: Van der Ploeg et al., 2024

The third example is a longitudinal oral microbiome gingivitis intervention study [7]. Full details on the microbiome sampling and pre-processing are described in the original publication. We analyse only one oral microbiome niche, the supragingival plaque at the upper jaw lingual surfaces, which was measured in 41 subjects at 7 time points: day -14 (baseline), 0 (baseline, start of intervention), 2, 5, 9, 14 (end of intervention) and 21 (resolution). Using quantitative light-induced fluorescence [31–33], the subjects were stratified into low, medium and high responders based on the amount of red fluorescent plaque on day 14.

The ASV data and taxonomic information were processed using the `parafac4microbiome` package (v1.0.0) in R (v4.3.2; [29]). The supplied dataset is a three-way array of counts of size 41 subjects x 2253 microbial (ASV) abundances x 7 time points. In accordance with the original analysis, ASVs were kept if the sparsity was ≤50% in any response group. This step resulted in 65 of the 2.253 ASVs being selected. Subsequently, a centred-log ratio transformation using a pseudo-count of 1 was performed to correct for compositionality [22, 23]. The data was then centred across the subject mode and scaled within the feature mode. The selected PARAFAC model in this study yields a slightly different clustering result from the original analysis due to differences in the random initialization step between Matlab and R. In the original analysis, Matlab was used, whereas in this study we used the `parafac4microbiome` package in R.

251

252 *Selecting the appropriate number of PARAFAC components*

253 Like in Principal Component Analysis, the number of biologically meaningful components that can
254 be extracted from longitudinal microbiome data is limited. After a given number of components,
255 PARAFAC will start to model noise. As such, one challenge of PARAFAC modelling for longitudinal
256 microbiome data is selecting the appropriate number of components. A model with the correct
257 number of components will (1) converge quickly, (2) not have highly similar loading vectors in any
258 mode, (3) describe meaningful biological phenomena and (4) describe these phenomena even when
259 one or more samples are removed from the data.

260 Various diagnostics and methods have been proposed and need to be manually examined to
261 determine the appropriate number of components for a PARAFAC model [34]. We propose a two-
262 step approach for selecting the correct number of components. First, we investigate quick
263 convergence and similarity of the loading vectors by randomly initializing a number of PARAFAC
264 models for 1-5 components. Next, we investigate the stability of the model using a jack-knife
265 procedure while keeping the initialized vectors the same using a singular value decomposition. In
266 the supplied R package, this procedure is automatically performed and visualized for inspection by
267 the user with the functions assessNumComponents() and modelStabilityCheck(), respectively. All
268 output plots examined to select the number of components for each dataset are shown in
269 **Supplementary Figures**. The metrics of interest are summarized below.

270 *Metrics for assessing the appropriate number of components*

271 For each dataset, we have investigated PARAFAC models with 1-5 components and a number of
272 random initializations, depending on the size of the data and how quickly the models converge. In
273 the supplied R package, this procedure is automatically performed and visualized for inspection by
274 the user with the function assessNumComponents(). We examine the following metrics of interest.

275 In PCA, the residual sum-of-squares are plotted against the number of components in a scree-plot
276 [35, 36]. The appropriate number of components is then the area of the plot where the residual sum-
277 of-squares flattens out. For PARAFAC this principle holds as well [18].

278 The alternating least-squares implementation of PARAFAC that is used throughout this paper
279 iteratively updates the loading vectors until the convergence threshold is reached. Therefore, the

280 number of iterations needed to converge is a useful metric to inspect for the correct number of
281 components. Randomly initialized PARAFAC models with the correct number of components
282 should converge quickly compared to models with an extra component modelling noise.

283 The core consistency diagnostic (CORCONDIA) indicates how well the assumed tri-linearity of the
284 PARAFAC model holds [37]. The range of CORCONDIA is typically 0-100 but can be negative. The
285 number of components in PARAFAC models with a CORCONDIA of 50-100 can be considered
286 appropriate [18, 37]. CORCONDIA in the `parafac4microbiome` is calculated using an
287 implementation from the multiway package that was modified to be robust towards missing values
288 (version 1.0-6; [19]).

289 The components of a PARAFAC model are non-orthogonal by default, which can sometimes lead
290 to very similar loading vectors if the model cannot be estimated properly. The Tucker congruence
291 coefficient identifies this similarity between pairs of loading vectors of the same mode [38, 39]. A
292 value in the range of 0.85-0.94 suggests that the loading vectors are fairly similar, whereas a value
293 of 0.95-1 means that the loading vectors are essentially equal [38]. The Tucker congruence
294 coefficient in the `parafac4microbiome` package uses the implementation from the multiway R
295 package (version 1.0-6; [19]).

296 *Assessing model stability through the jack-knifing of samples*

297 We expect a PARAFAC model with the correct number of components to describe meaningful
298 biological phenomena, even when one or several samples are removed from the data. We test this
299 by performing iterative jack-knifing of samples [40]. In the supplied R package, this procedure is
300 automatically performed and visualized for inspection by the user with the function
301 modelStabilityCheck(). With this jack-knifing approach, one random sample is removed in every
302 iteration, after which a PARAFAC model is fitted to the data. We perform balanced jack-knifing,
303 removing one sample from every subject group in every iteration, if such subject metadata is
304 available. If this is impossible due to unequally sized groups, samples are randomly chosen. We
305 choose to initialise the modelling algorithm using a singular value decomposition for every iteration
306 to ensure that instability can only be caused by the removed sample(s). A PARAFAC model with
307 the correct number of components will be stable (i.e. have a small standard deviation for every
308 loading) across all jack-knifing iterations.

309

310

311 *Post-hoc clustering of PARAFAC loadings*

312 Clustering of the feature loadings can help identify groups of microbiota that are connected to a
313 particular time profile in the PARAFAC model. We performed the clustering procedure of the
314 feature loadings as follows. In accordance with the original analysis in van der Ploeg et al. [7],
315 features were not considered if the variation explained by the model for their entire time trajectory
316 was lower than the average or the congruence loading of the feature was lower than 0.4 [41]. The
317 remaining features were then clustered based on their modelled time trajectories using the K-
318 medoids algorithm from the cluster R package (v2.1.4; [42]) with 50 random starts to be robust
319 against outliers. The number of clusters was determined using the within-cluster sum of squares,
320 silhouette width and gap statistic metrics as reported by the factoextra R package (**Supplementary**
321 **Figure 15**; v1.0.7; [43]).

322 **Results**

323 *PARAFAC can identify the main time-resolved variation in longitudinal microbiome data*

324 The first example is a long time series of experimental in vitro microbiomes [6]. We analyse only
325 one experimental set-up, the aquatic microbiome culture in a peptone medium, which was
326 performed in eight replicates and measured daily for 110 days. With this example, we aim to
327 visualize the main time-resolved variation that can be modelled through PARAFAC.

328 We inspected the unprocessed count data by computing the relative abundances for every replicate
329 culture at every time point (**Figure 2A**). In most samples, 2 amplicon sequence variants (ASVs)
330 belonging to *Aeromonas* spp. comprised most of the counts. The exceptions were: the emergence of
331 *Cedecea*, *Burkholderia-Caballeronia-Paraburkholderia*, and *Pandoraea* species between days 50-110
332 in replicate culture 4, the emergence of *Clostridioides mangenotii* and a *Lachnoclostridium* species
333 between days 80-100 in replicate cultures 4 and 8, the disappearance of a *Salmonella*, *Propionivibrio*,
334 *Microbacter* and *Herbaspirillum* species between days 0-60 in replicate culture 3, the transient
335 appearance of a *Microbacter* species in replicate cultures 2 and 8, and a (small) increase in a
336 *Salmonella* species between days 100-110 in most replicate cultures. There was no clear time profile
337 visible that all replicates had in common.

338 We inspected PARAFAC modelling performance and stability for the processed count data using 1-
339 5 components. The core consistency diagnostic (CORCONDIA) scores dropped below 80 for 5
340 components but were comparable between 3 and 4 components. The three and four-component
341 models explained 37% and 42% on average, respectively (**Supplementary Figure 2**). Performing a
342 jack-knifing procedure showed that the three-component PARAFAC models were more stable than
343 the four-component models (**Supplementary Figures 3-4**). Hence, we selected the model explaining
344 the most variation from all randomly initialized models of three components. The selected model
345 explained 39.0% of the variation (**Figure 2B**).

346 As expected, the selected PARAFAC model did not describe *Aeromonas* ASVs, because these do not
347 vary strongly over time. The first component described the emergence of *Cedecea*, *Burkholderia-*
348 *Caballeronia-Paraburkholderia*, and *Pandoraea* species between days 50-110, which was unique to
349 replicate culture 4. The second component described the emergence of *Clostridioides mangenotii*
350 and a *Lachnoclostridium* species between days 80-100 in replicate cultures 4 and 8. The third
351 component described the disappearance of a *Salmonella*, *Propionivibrio*, *Microbacter,* and
352 *Herbaspirillum* species between days 0-60, which was unique to replicate culture 3.

353 While this example highlights the strength of PARAFAC: identifying and modelling the largest
354 sources of variation in a longitudinal microbiome dataset, the PARAFAC model only described the
355 variation partially (39%). As a result, the transient appearance of a *Microbacter* species in replicate
356 cultures 2 and 8, nor the (small) increase in a *Salmonella* species between days 100-110 in most
357 replicate cultures, were not modelled.

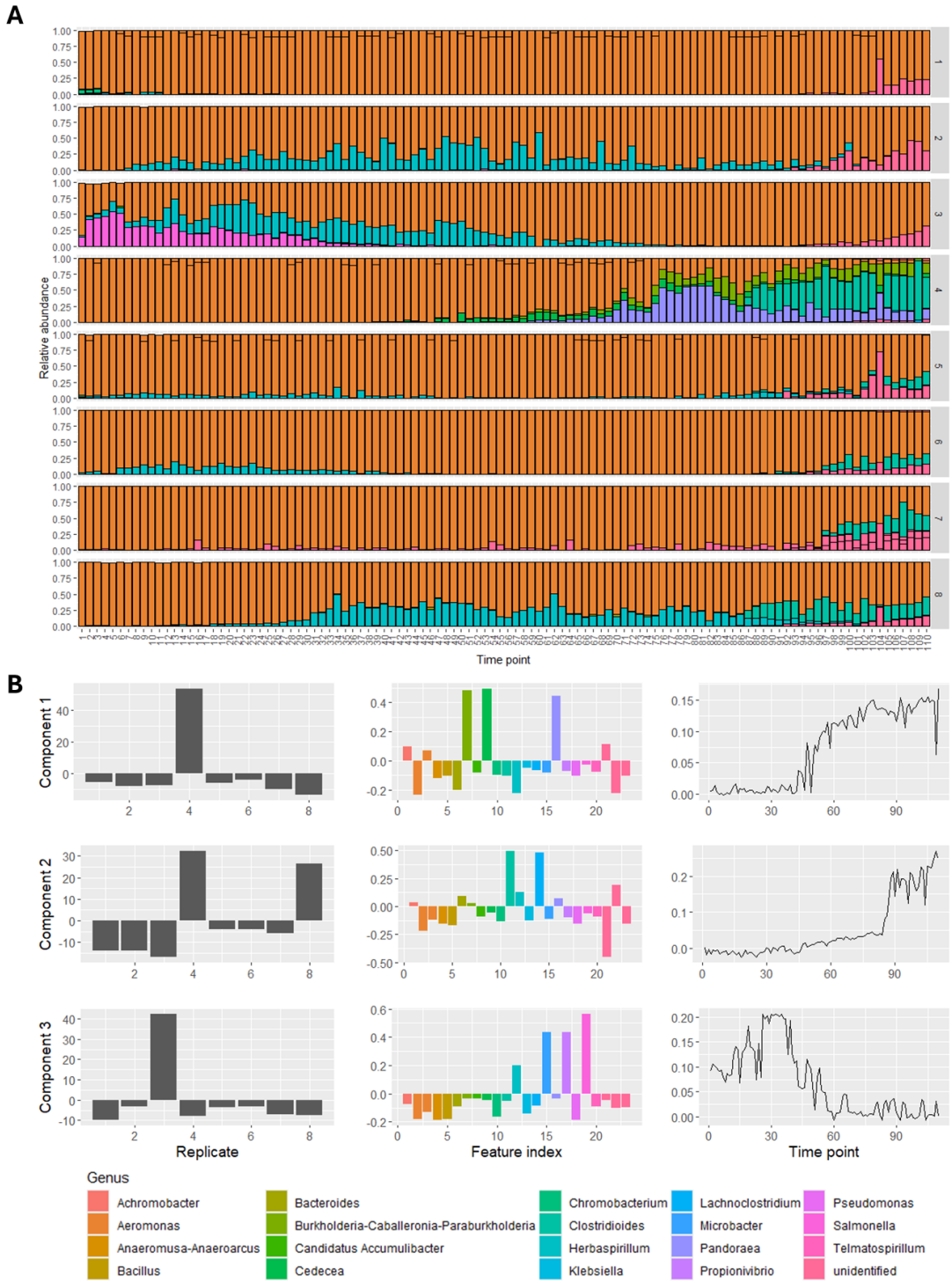


Figure 2: Overview of the Fujita et al., 2023 dataset. (A) Relative abundance plot of the dataset, split per replicate. In most samples, Aeromonas species comprise the vast majority of the counts. In the latter half of the samples in replicate 4 and the first half of the samples in replicate 3 other species are present that comprise the majority. There is no clear time profile visible that all replicates have in common. (B) The best three-component PARAFAC model of the processed data describing 39.0% of the variation. The components are shown in the rows, while the replicate, microbial abundance and time modes are shown in the columns. The first component describes the emergence of a Cedecea, Burkholderia-Caballeronia-Paraburkholderia and Pandoraea species between days 50-110, which is unique to replicate 4. The second component describes the emergence of Clostridioides mangenotii and a Lachnoclostridium species between days 80-100 in replicates 4 and 8. The third component describes the disappearance of a Salmonella, Propionivibrio, Microbacter and Herbaspirillum species between days 0-60, which is unique to replicate 3.

369 *PARAFAC finds differences between subject groups and enhances comparative analyses*

370 The second example is a longitudinal infant faecal microbiome study [8]. Samples were collected
371 from 314 vaginally born and 282 caesarean-section born infants at four time points: 4, 7 and 21 days
372 after birth and a follow-up sample during infancy (between 4-12 months after birth). One of the
373 four measurements is missing for most infants in this cohort, and we keep the missing data in the
374 data cube for the PARAFAC algorithm to impute (**Supplementary Table 1-2**). With this example,
375 we show how PARAFAC can identify the differences between the subject groups and enhance the
376 comparative analysis done by the original authors despite a moderate amount of missing data.

377 The authors stated that there was a very clear distinction in the microbiomes of vaginally and
378 caesarean-section born children, also over time. To analyse this distinction, we inspected the
379 unprocessed count data by computing the average relative abundance for the vaginal and caesarean
380 groups at every time point (**Figure 3A**). This confirmed a clear difference at the phylum level
381 between the vaginal and caesarean groups at day 4, 7 and 21, which disappeared in infancy.

382 We inspected PARAFAC modelling performance and stability for the processed count data using 1-
383 5 components. The CORCONDIA scores dropped below 80 at 5 components and the 3-5 component
384 models did not converge within 500 iterations (**Supplementary Figure 6**). While the time profiles in
385 the two-component models look similar, the Tucker Congruence Coefficient plots show that they
386 are not the same (**Supplementary Figure 7**). Performing a jack-knifing procedure showed that the
387 two-component model was more stable than the three-component model (**Supplementary figures**
388 **8-9**). Hence, we selected the model explaining the most variation from all randomly initialized
389 models of two components. The selected model explained 11.18% of the variation (**Figure 3B**).

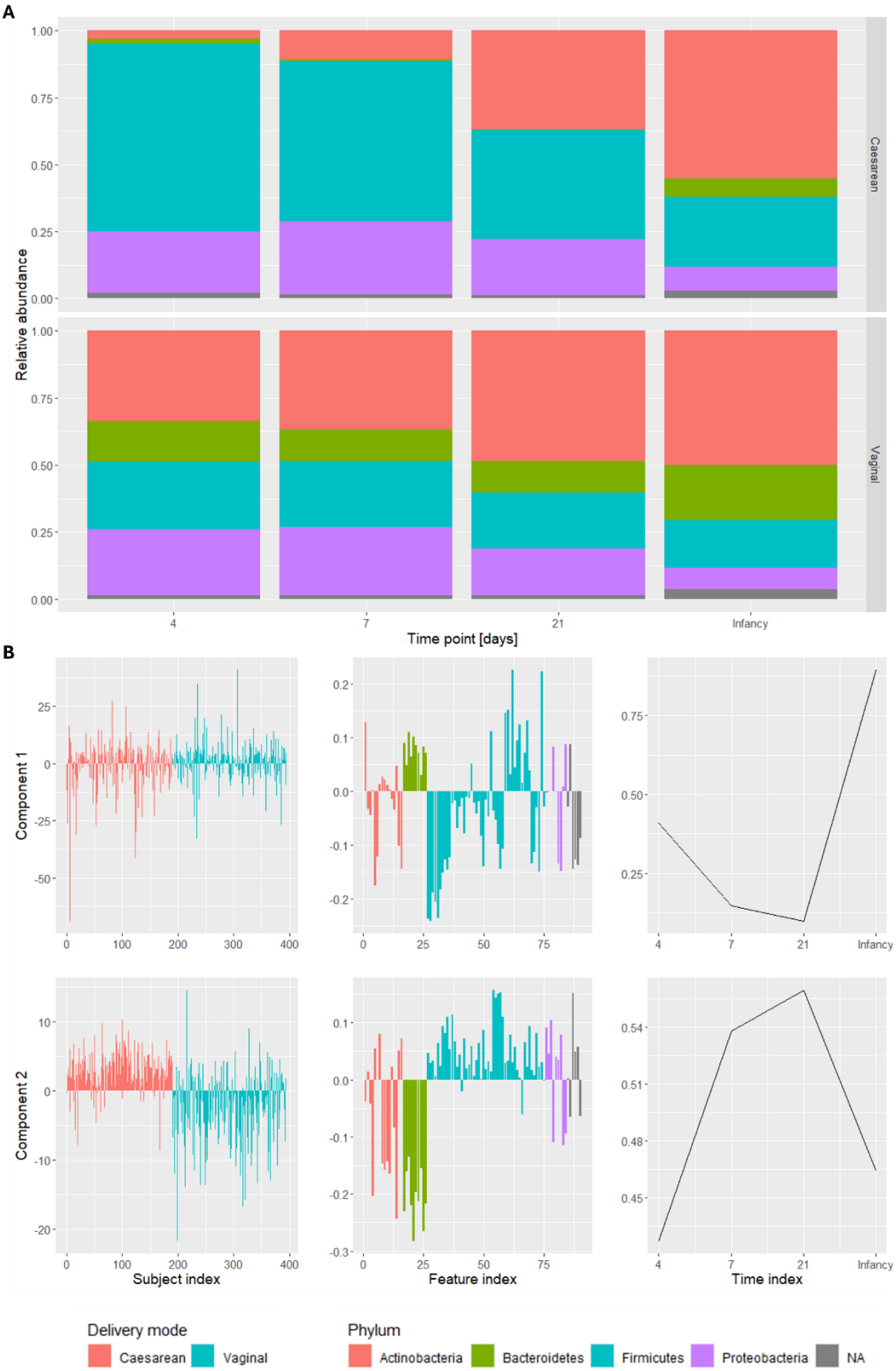


Figure 3: Overview of the Shao et al., 2019 dataset. (A) Overview of the average relative abundance per time point in vaginal and caesarean-section born infants. The birth mode groups are distinguishable between time points 4 and 21 and become similar at infancy. (B) The best two-component PARAFAC model of the processed data, explaining 11.18% of the variation. Component one mainly describes differences in feeding method. Component two describes the difference in microbiomes between the birth mode groups.

396 The first component of the selected model described the antibiotics given to the baby in the hospital
397 (p=4.3e-08; Benjamini-Hochberg corrected Wilcoxon rank-sum test; **Supplementary Table 3**), the
398 feeding mode (p=4e-05) and the *Bacteroides* profile described by the original authors (p=0.024). The
399 second component modelled birth mode (p=1.1e-15) and the *Bacteroides* profile (p=1.3e-15).

400 Inspection of the positive loadings in the first component revealed that PARAFAC described the
401 genera that were enriched in formula fed children: *Flavonifractor plautii, Blautia obeum/wexlerae,*
402 *Tyzzerella nexilis, Eggerthella lenta, Lachnoclostridium* sp. and *Bacteroides* sp. It also described the
403 genera that were enriched in breast fed children with negative loadings: *Streptococcus (peroris,*
404 *pneumoniae, oralis, infantis, parasanguinis, australis), Actinobacteria* sp., *Haemophilus*
405 *influenzae/parainfluenzae, Rothia mucilaginosa, Veilonella (dispar, atypica), Gemella (haemolysans,*
406 *sanguinis)* and *Actinomyces* sp. While not described in the original analysis, these results partially
407 agree with other literature [44–46].

408 Inspection of the positive loadings in the second component revealed that PARAFAC described the
409 genera that were enriched in caesarean-section born infants and depleted in vaginally born infants:
410 *Clostridium perfringens, Veillonella (dispar, atypica, rogosae, parvula), Enterococcus faecalis,*
411 *Streptococcus parasanguinis,* and *Klebsiella michiganensis/oxytoca*. This result is consistent with
412 the original analysis and other studies investigating the development of the infant gut microbiome
413 in hospital environments [47, 48].

414 Likewise, the negatively loaded genera were enriched in vaginally born infants and depleted in
415 caesarean-section born infants: *Bacteroides (dorei/vulgatus, uniformis, faecis/thetaiotaomicron,*
416 *caecimuris, fragilis), Parabacteroides (distasonis, merdae, caccae, stercoris), Collinsella aerofaciens,*
417 *Bifidobacterium (longum, adolescentis, catenulatum/kashiwanohense, pseudocatenulatum,*
418 *ruminantium), Sutterella wadsworthensis,* and *Escherichia coli*. This result is consistent with the
419 original analysis and with reports from other studies [49, 50]. The exception is *Sutterella*
420 *wadsworthensis*, which is not commonly reported to be enriched.

421 Taken together, the PARAFAC model of this example dataset confirmed previously found
422 differences between the caesarean-section and vaginally born infants. It also found novel genera
423 that describe the differences between the birth mode groups and feeding mode groups. Additionally,
424 the large amount of missing data did not hamper the modelling procedure. In summary, the

425 PARAFAC model of this dataset confirms and enhances the results from the comparative analysis
426 done by the original authors.

427 *Post-hoc clustering of PARAFAC results can help identify microbial groups of interest*

428 The third example is a longitudinal oral microbiome gingivitis intervention study [7]. Here, we
429 analyse one oral microbiome niche, the supragingival plaque at the upper jaw lingual surfaces,
430 which was measured in 41 subjects at 7 time points: day -14 (baseline), 0 (baseline, start of
431 intervention), 2, 5, 9, 14 (end of intervention) and 21 (resolution). Using quantitative light-induced
432 fluorescence [31–33], the subjects were stratified into low, medium and high responders based on
433 the amount of red fluorescent dental plaque on day 14. With this example, we show how a post-hoc
434 clustering approach of a PARAFAC model can help identify groups of microbiota related to the
435 intervention.

436 We inspected PARAFAC modelling performance and stability for the processed count data using 1-
437 5 components. The CORCONDIA scores dropped below 80 for 4 and 5 components and the sum of
438 squared errors began to flatten between 2-3 components (**Supplementary Figure 11**). The three-
439 component models had very similar feature and time mode loadings between several components,
440 making them not suitable for describing the data (**Supplementary Figure 12**). This was confirmed
441 by the jack-knifing procedure (**Supplementary Figures 13-14**). Hence, we selected the model
442 explaining the most variation from all randomly initialized models of two components. The selected
443 model explained 19.21% of the variation (**Figure 4A**).

444

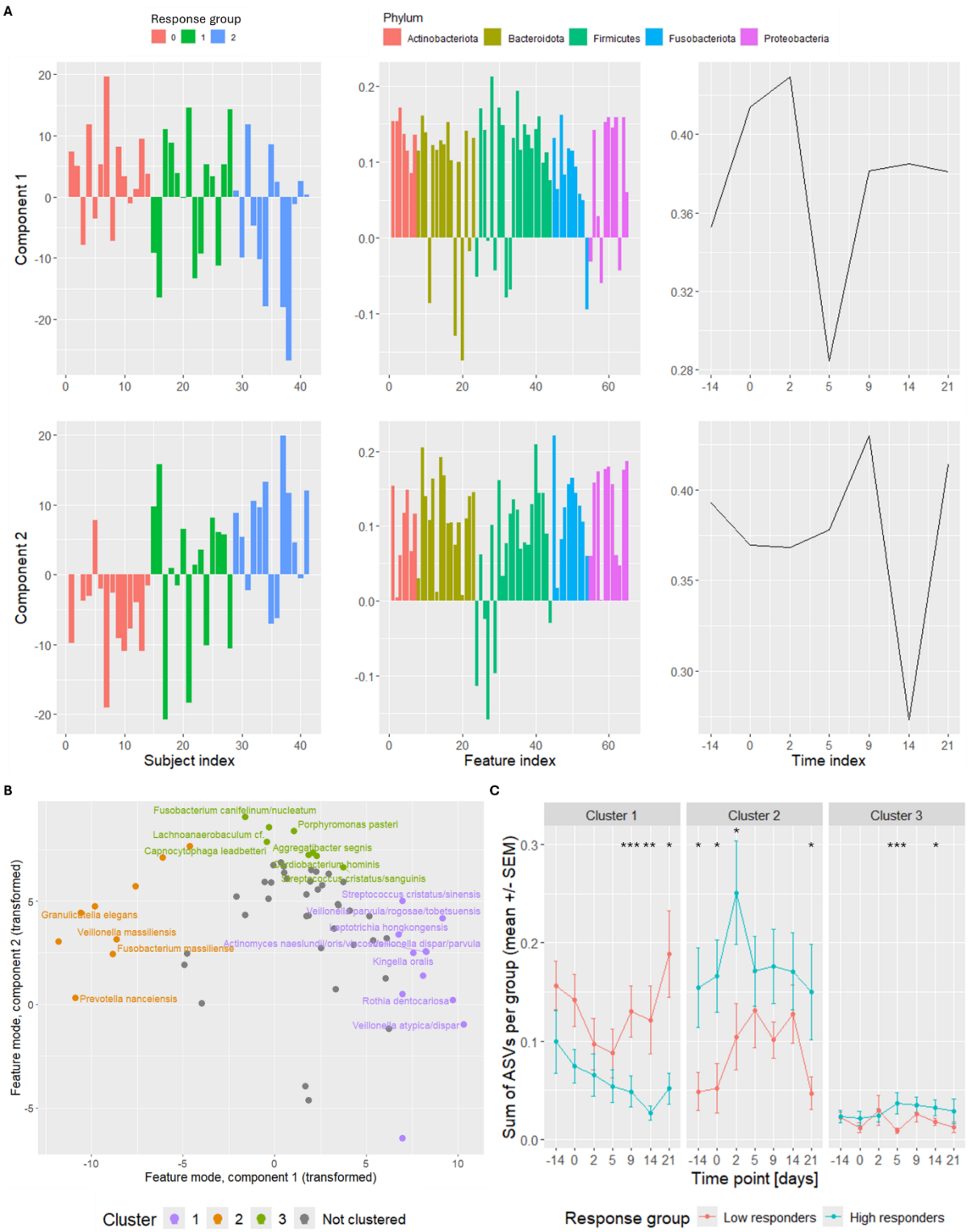


Figure 4: Overview of the van der Ploeg et al., 2024 dataset. (A) The best two-component PARAFAC model of the processed data, explaining 19.21% of the variation. Both components describe the amount of red fluorescent plaque over time and the percentage of sites that are covered in plaque. (B) Scatter plot showing the transformed feature loadings (Supplementary Methods) of PARAFAC component 1 versus component 2. Features were clustered into three groups after filtering. Only amplicon sequence variants (ASVs) with taxonomic annotations resolved to species level are labelled. (C) Overview of the relative abundance sum per ASV cluster in the low and high response groups. Error bars correspond to the standard error of the mean (SEM) across all samples per response group. The mean difference between the high and low response groups per time point was tested using a Benjamini-Hochberg corrected permutation test of 999 iterations (: p≤0.05; **: p≤0.01; ***: p≤0.001). Mid responders are not shown for visual clarity. See* ***Table 2*** *for the list of identified species per ASV cluster.*

455 To interpret the models, we tested the correlation of the subject loadings with the measured
456 parameters of gingivitis. This approach revealed that both components of the PARAFAC model
457 significantly described the amount of red fluorescent plaque over time (p=8.5e-6 and p=4.2e-7,
458 respectively; Benjamini-Hochberg corrected Pearson correlation test; **Supplementary Table 4**) and the
459 percentage of sites that were covered by plaque (p=0.024 and p=0.002, respectively).

460 We identified microbial subcommunities with common responses by selecting and then clustering
461 well-modelled amplicon sequence variants (ASVs) based on their fitted response (**Figure 4B, Table 2**).
462 The sum of the relative abundances of the ASVs in each cluster was then determined to test the
463 difference between the microbiomes of individuals in the low and high response groups at every time
464 point. This approach revealed that the relative abundance sum of ASV cluster 1 was significantly higher
465 in the low responder individuals compared to the high responder individuals on days 9, 14 and 21
466 (**Figure 4C, Supplementary Table 5**; p≤0.05; Benjamini-Hochberg corrected permutation test of mean
467 difference). Similarly, the relative abundance sum of ASV cluster 2 was found to be significantly higher
468 in high responder individuals compared to low responder individuals on days -14, 0, 2 and 21. The
469 relative abundance sum of ASV cluster 3 was significantly higher in high responder individuals
470 compared to low responder individuals on days 5, and 21. Taken together, the PARAFAC models of
471 plaque microbiomes described microbial subcommunities with common dynamics that connect to the
472 magnitude of the gingivitis response.

473 *Table 2: Overview of all identified species per amplicon sequence variant (ASV) cluster. ASVs within the same cluster show a*
474 *similar response over time to the gingivitis intervention. Only clustered ASVs with taxonomic annotations resolved to species*
475 *level are shown.*

Cluster number	Number of members	Identified species
1	11	*Actinomyces naeslundii/oris/viscosus, Kingella oralis, Leptotrichia honghongensis, Rothia dentocariosa, Streptococcus cristatus/sinensis, Veillonella atypica/dispar, Veillonella dispar/parvula, Veillonella parvula/rogosae/tobetsuensis*
2	9	*Fusobacterium massiliense, Granulicatella elegans, Prevotella nanceiensis, Veillonella massiliensis*
3	8	*Aggregatibacter segnis, Capnocytophaga leadbetteri, Cardiobacterium hominis, Fusobacterium canifelinum/nucleatum, Lachnoanaerobaculum* cf.*, Porphyromonas pasteri, Streptococcus cristatus/sanguinis*

476

477

478

Discussion

Using three example studies [6–8], we show that the unsupervised modelling approach of Parallel Factor Analysis (PARAFAC) can (1) identify the main time-resolved variation in longitudinal microbiome data, (2) find differences between subject groups and enhance comparative analyses and (3) helps identify feature groups of interest using a post-hoc clustering approach. In addition, we show that PARAFAC is robust to missing measurements in the Shao et al., 2019 [8] example. We also show in the Shao et al., 2019 [8] and van der Ploeg et al., 2024 [7] examples that the modelled variation is related to clinically relevant parameters. However, many other multi-way approaches exist that are more flexible in their modelling procedure [28]. For example, Tucker3 allows for interactions between the components [51]. PARAFAC2 allows for time shifts that are unique per subject [18, 52]. Multi-way regression, in which the PARAFAC components are supervised using a response variable, is also possible with N-way partial least squares [53]. As highlighted in the Fujita et al. example study [6], these multi-way dimension reduction methods require commonality between the subjects to be useful. Further research is needed to investigate the relevance of these methods in the context of longitudinal microbiome data analysis.

In PARAFAC, the nature of the third mode in the multi-way array is not limited to time. For example, a recent study has shown that soil treatment can be used as the third mode to identify relationships between the abundances of mycorrhizal fungi and plant diversity and history [54]. Additionally, the principles shown in this work are generalizable to four or more modes if that better represents the data that was generated by the study design.

In this work we have not elaborated on the various options that exist for processing microbial abundance data. For example, in a related paper on tensor factorization of longitudinal microbiome data [55] the robust-centred log-ratio was suggested as an approach to transform the data and to correct for compositionality. This approach instead calculates the geometric mean by ignoring the zeroes and setting them to missing values after the transformation. This makes all zeroes available for imputation by the PARAFAC model, but ignores the different nature that zeroes can potentially have [24]. Hence, we decided to use a centred log-ratio (CLR) transformation with a pseudo-count of 1 as this keeps the zero values intact. Further research is needed to investigate the appropriateness of the various methods for (longitudinal) microbiome count data in describing biological phenomena of interest.

Conclusions

We have shown that Parallel Factor Analysis is an applicable method for longitudinal microbiome data analysis across a wide range of microbial environments. Our results demonstrate that PARAFAC is able to (1) model the main time-resolved variation in longitudinal microbiome data, (2) find differences between subject groups and enhance comparative analyses despite a moderate amount of missing data, and (3) help identify feature groups of interest. The analyses and the example datasets with the resulting figures are implemented in the R package `parafac4microbiome`, which is available online at https://github.com/GRvanderPloeg/parafac4microbiome.

List of abbreviations

- ALS: Alternating Least-Squares
- ASV: Amplicon Sequence Variant
- CLR: Centred Log-Ratio
- CORCONDIA: Core Consistency Diagnostic
- mOTU: Marker-gene based Operational Taxonomic Unit
- PCA: Principal Component Analysis
- PARAFAC: Parallel Factor Analysis

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of the data and materials

The package used during the current study is available in the GitHub repository, https://doi.org/10.5281/zenodo.11066742. The datasets generated and/or analysed during the current study are available in the GitHub repository, https://doi.org/10.5281/zenodo.11072652.

Competing interests

The authors declare that they have no competing interests.

Funding

GRvdP was funded by a grant from the University of Amsterdam, Research Priority Area on Personal Microbiome Health.

Authors' contributions

G.R. van der Ploeg: conceptualization, visualization, formal analysis, methodology, programming, writing – original draft; J.A. Westerhuis: conceptualization, methodology, supervision, formal analysis, writing – original draft; A. Heintz-Buschart: conceptualization, formal analysis, supervision, project administration, writing – original draft; A.K. Smilde: conceptualization, methodology, supervision, funding acquisition, writing – review & editing.

Acknowledgements

We want to thank Jesse Alderliesten (UU/UvA) for their useful discussions and suggestions.

References

1. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature. 2019;569:655–62.

2. Pascal M, Perez-Gordo M, Caballero T, Escribese MM, Lopez Longo MN, Luengo O, et al. Microbiome and Allergic Diseases. Front Immunol. 2018;9.

3. Saus E, Iraola-Guzmán S, Willis JR, Brunet-Vega A, Gabaldón T. Microbiome and colorectal cancer: Roles in carcinogenesis and clinical potential. Mol Aspects Med. 2019;69:93–106.

4. Gerber GK. Longitudinal Microbiome Data Analysis. In: Metagenomics for Microbiology. Elsevier; 2015. p. 97–111.

5. Schmidt TSB, Raes J, Bork P. The Human Gut Microbiome: From Association to Modulation. Cell. 2018;172:1198–215.

6. Fujita H, Ushio M, Suzuki K, Abe MS, Yamamichi M, Iwayama K, et al. Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. Microbiome. 2023;11:63.

7. van der Ploeg GR, Brandt B, Keijser B, Veen M van der, Volgenant C, Zaura E, et al. Multi-way modelling of oral microbial dynamics and host-microbiome interactions during induced gingivitis. 2024; doi: 10.1101/2024.03.18.585469.

8. Shao Y, Forster SC, Tsaliki E, Vervier K, Strang A, Simpson N, et al. Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. Nature. 2019;574:117–21.

9. Diggle P. Analysis of longitudinal data. Oxford university press; 2002.

10. Fitzmaurice GM, Laird NM, Ware JH. Applied longitudinal analysis. John Wiley & Sons; 2012.

11. Kodikara S, Ellul S, Lê Cao K-A. Statistical challenges in longitudinal microbiome data analysis. Brief Bioinform. 2022;23:bbac273.

12. Bro R, Smilde AK. Principal component analysis. Anal Methods. 2014;6:2812–31.

13. Jolliffe I. Principal Component Analysis. In: Wiley StatsRef: Statistics Reference Online. John Wiley & Sons, Ltd; 2014.

14. Carroll JD, Chang J-J. Analysis of individual differences in multidimensional scaling via an N-way generalization of “Eckart-Young” decomposition. Psychometrika. 1970;35:283–319.

15. Harshman RA. Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multimodal factor analysis. 1970.

16. Tomasi G, Bro R. A comparison of algorithms for fitting the PARAFAC model. Comput Stat Data Anal. 2006;50:1700–34.

17. Yates F. The analysis of replicated experiments when the field results are incomplete. Emp J Exp Agric. 1933;1:129–42.

18. Bro R. Multi-way analysis in the food industry. Models Algorithms Appl Acad Proefschr Dinam. 1998.

19. Helwig NE, Helwig MNE. Package ‘multiway.’ 2019.

20. Kruskal JB. More factors than subjects, tests and treatments: An indeterminacy theorem for canonical decomposition and individual differences scaling. Psychometrika. 1976;41:281–93.

21. Kruskal JB. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. Linear Algebra Its Appl. 1977;18:95–138.

22. Aitchison J. The Statistical Analysis of Compositional Data. J R Stat Soc Ser B Methodol. 1982;44:139–60.

23. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. Front Microbiol. 2017;8:2224.

24. Silverman JD, Roche K, Mukherjee S, David LA. Naught all zeros in sequence count data are the same. Comput Struct Biotechnol J. 2020;18:2789–98.

25. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. Nat Methods. 2013;10:1200–2.

26. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol. 2011;12:R60.

27. Bro R, Smilde AK. Centering and scaling in component analysis. J Chemom. 2003;17:16–33.

28. Smilde AK, Bro R, Geladi P. Multi-way analysis: applications in the chemical sciences. John Wiley & Sons; 2005.

29. R Core Team R. R: A language and environment for statistical computing. 2013.

601 30. Milanese A, Mende DR, Paoli L, Salazar G, Ruscheweyh H-J, Cuenca M, et al. Microbial abundance,
602 activity and population genomic profiling with mOTUs2. Nat Commun. 2019;10:1014.

603 31. Heinrich-Weltzien R, Kühnisch J, Van Der Veen M, De Josselin De Jong E, Stößer L. Quantitative
604 light-induced fluorescence (QLF) - A potential method for the dental practitioner. Quintessence Int.
605 2003;34:181–8.

606 32. van der Veen MH, Volgenant CMC, Keijser B, ten Cate J (Bob) M, Crielaard W. Dynamics of red
607 fluorescent dental plaque during experimental gingivitis—A cohort study. J Dent. 2016;48:71–6.

608 33. Volgenant CMC, Zaura E, Brandt BW, Buijs MJ, Tellez M, Malik G, et al. Red fluorescence of dental
609 plaque in children —A cross-sectional study. J Dent. 2017;58:40–7.

610 34. Bro R. PARAFAC. Tutorial and applications. Chemom Intell Lab Syst. 1997;:23.

611 35. Cattell RB. The Scree Test For The Number Of Factors. Multivar Behav Res. 1966;1:245–76.

612 36. Horn JL, Engstrom R. Cattell's Scree Test In Relation To Bartlett's Chi-Square Test And Other
613 Observations On The Number Of Factors Problem. Multivar Behav Res. 1979;14:283–300.

614 37. Bro R, Kiers H. A new Efficient Method for Determining the Number of Components in PARAFAC
615 Models. J Chemom. 2003;17:274–86.

616 38. Lorenzo-Seva U, ten Berge JMF. Tucker's congruence coefficient as a meaningful index of factor
617 similarity. Methodol Eur J Res Methods Behav Soc Sci. 2006;2:57–64.

618 39. Tucker LR. A method for synthesis of factor analysis studies. Educational Testing Service Princeton,
619 NJ; 1951.

620 40. Riu J, Bro R. Jack-knife technique for outlier detection and estimation of standard errors in PARAFAC
621 models. Chemom Intell Lab Syst. 2003;65:35–49.

622 41. Lorho G, Westad F, Bro R. Generalized correlation loadings: Extending correlation loadings to
623 congruence and to multi-way models. Chemom Intell Lab Syst. 2006;84:119–25.

624 42. Maechler M. Finding groups in data: Cluster analysis extended Rousseeuw et al. R Package Version.
625 2019;2:242–8.

626 43. Kassambara A, Mundt F. Package 'factoextra.' Extr Vis Results Multivar Data Anal. 2017;76.

627 44. Lee SA, Lim JY, Kim B-S, Cho SJ, Kim NY, Kim OB, et al. Comparison of the gut microbiota profile in
628 breast-fed and formula-fed Korean infants using pyrosequencing. Nutr Res Pract. 2015;9:242–8.

629 45. Ma J, Li Z, Zhang W, Zhang C, Zhang Y, Mei H, et al. Comparison of gut microbiota in exclusively
630 breast-fed and formula-fed babies: a study of 91 term infants. Sci Rep. 2020;10:15792.

631 46. Raspini B, Vacca M, Porri D, De Giuseppe R, Calabrese FM, Chieppa M, et al. Early Life Microbiota
632 Colonization at Six Months of Age: A Transitional Time Point. Front Cell Infect Microbiol. 2021;11.

633 47. Raveh-Sadka T, Firek B, Sharon I, Baker R, Brown CT, Thomas BC, et al. Evidence for persistent and
634 shared bacterial strains against a background of largely unique gut colonization in hospitalized
635 premature infants. ISME J. 2016;10:2817–30.

636 48. Stewart CJ, Skeath T, Nelson A, Fernstad SJ, Marrs EC, Perry JD, et al. Preterm gut microbiota and
637 metabolome following discharge from intensive care. Sci Rep. 2015;5:17141.

638 49. Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, et al. Dynamics and Stabilization
639 of the Human Gut Microbiome during the First Year of Life. Cell Host Microbe. 2015;17:690–703.

640 50. Wampach L, Heintz-Buschart A, Fritz JV, Ramiro-Garcia J, Habier J, Herold M, et al. Birth mode is
641 associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential.
642 Nat Commun. 2018;9:5091.

643 51. Kroonenberg PM, De Leeuw J. Principal component analysis of three-mode data by means of
644 alternating least squares algorithms. Psychometrika. 1980;45:69–97.

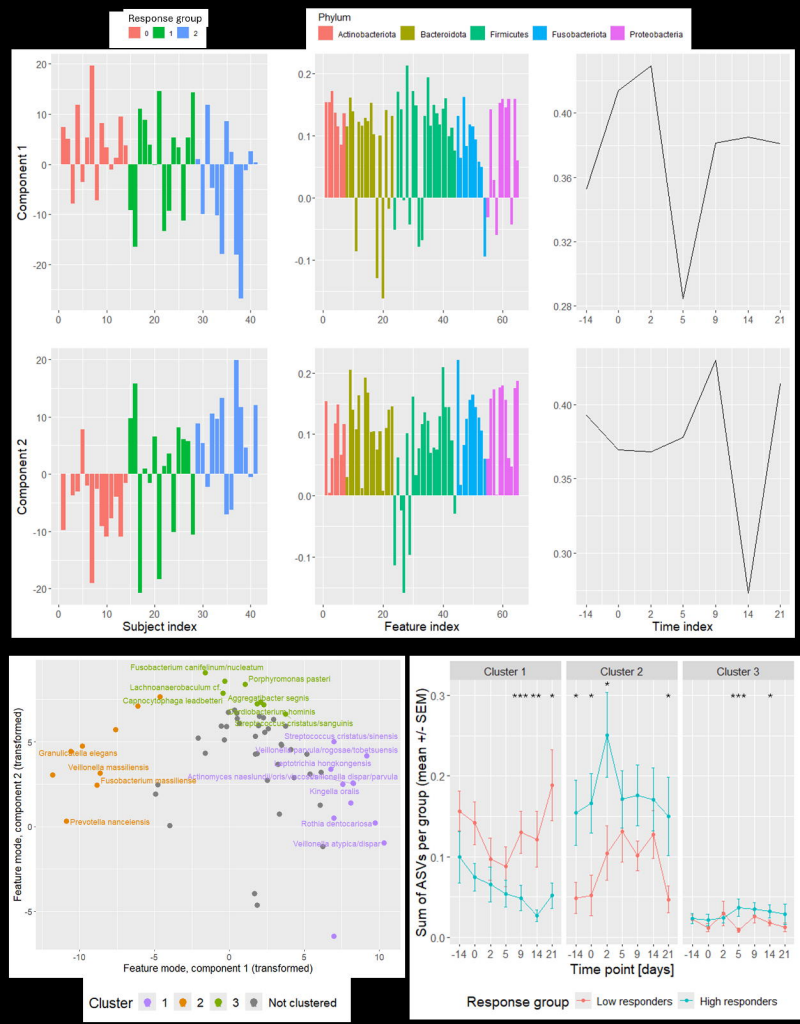
645 52. Kiers HAL, Ten Berge JMF, Bro R. PARAFAC2—Part I. A direct fitting algorithm for the PARAFAC2
646 model. J Chemom. 1999;13:275–94.

647 53. Bro R. Multiway calibration. Multilinear PLS. J Chemom. 1996;10:47–61.

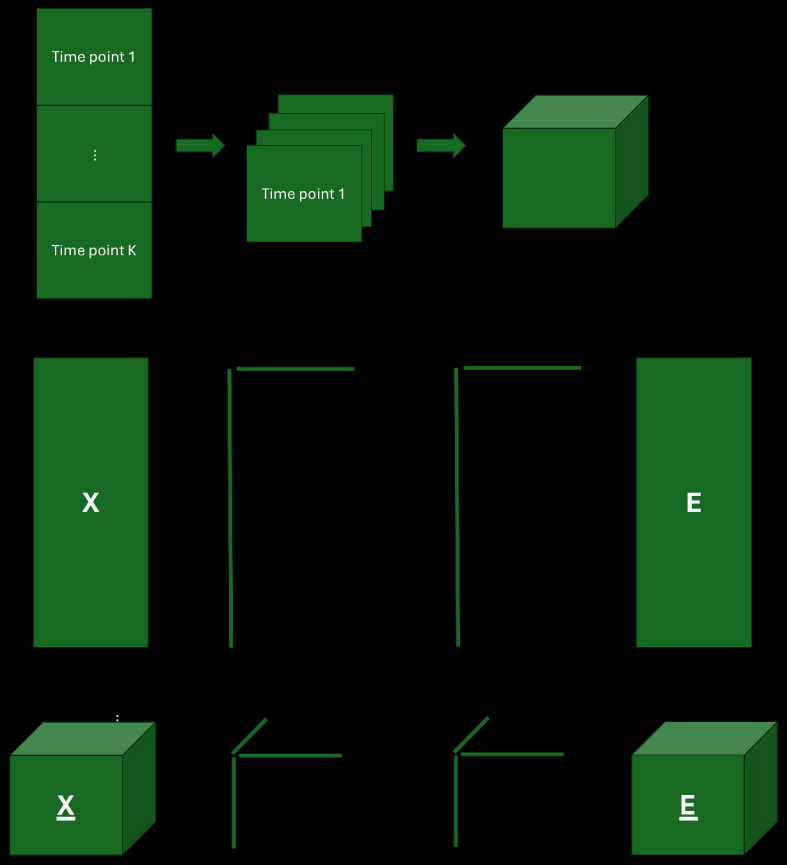
648 54. Albracht C, Solbach MD, Hennecke J, Bassi L, Van Der Ploeg GR, Eisenhauer N, et al. Common soil
649 history is more important than plant history for arbuscular mycorrhizal community assembly in an
650 experimental grassland diversity gradient. preprint. Microbiology; 2024.

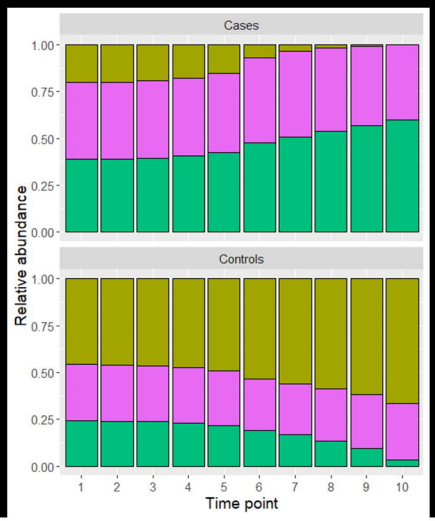
651 55. Martino C, Morton JT, Marotz CA, Thompson LR, Tripathi A, Knight R, et al. A Novel Sparse
652 Compositional Technique Reveals Microbial Perturbations. mSystems. 2019;4:e00016-19.

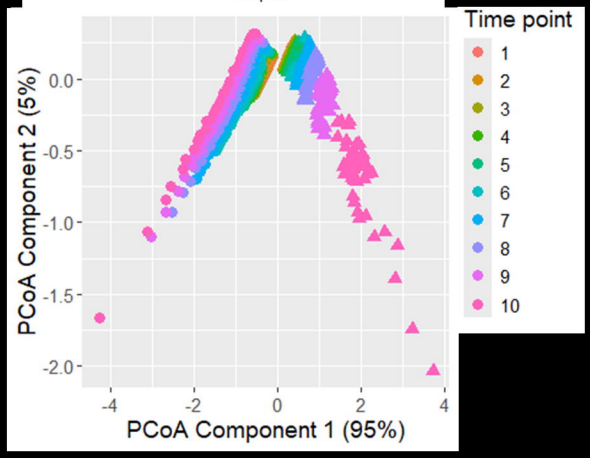
653

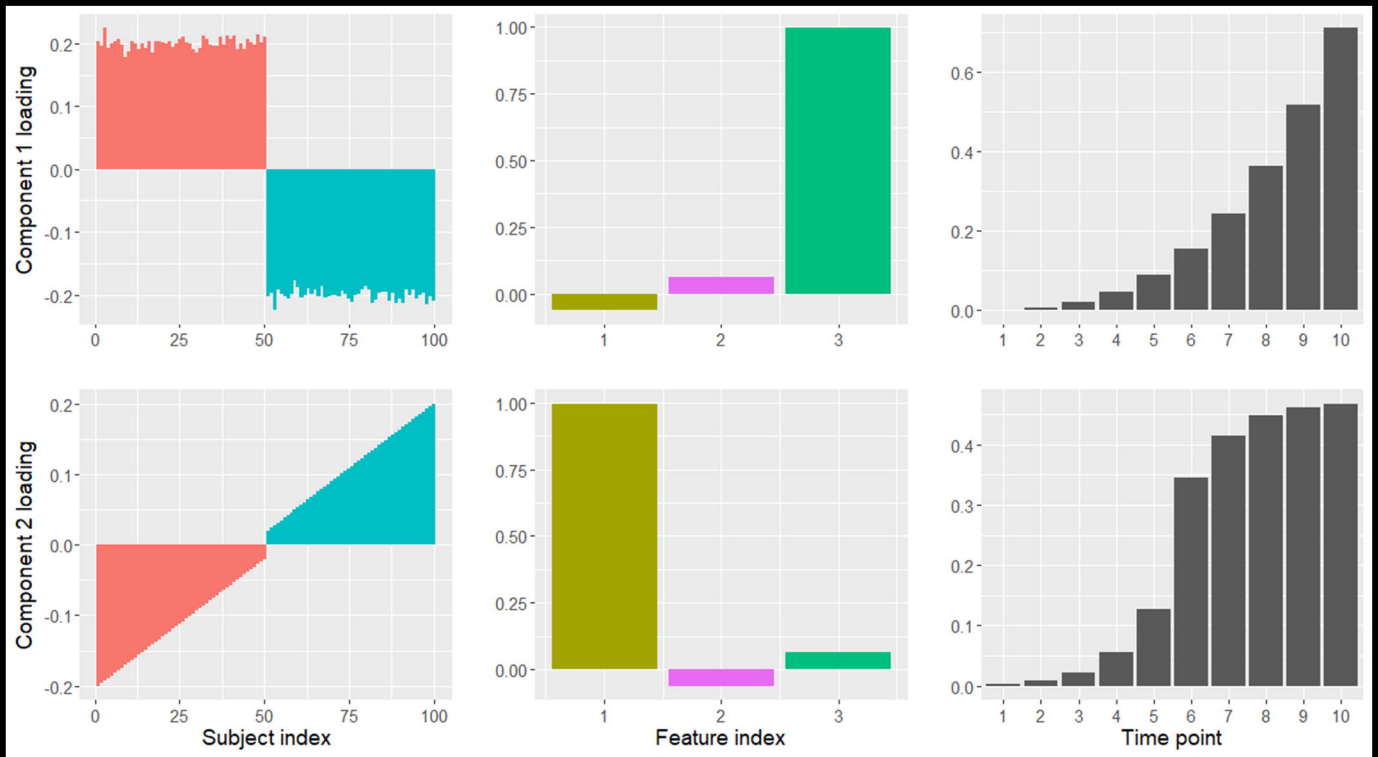


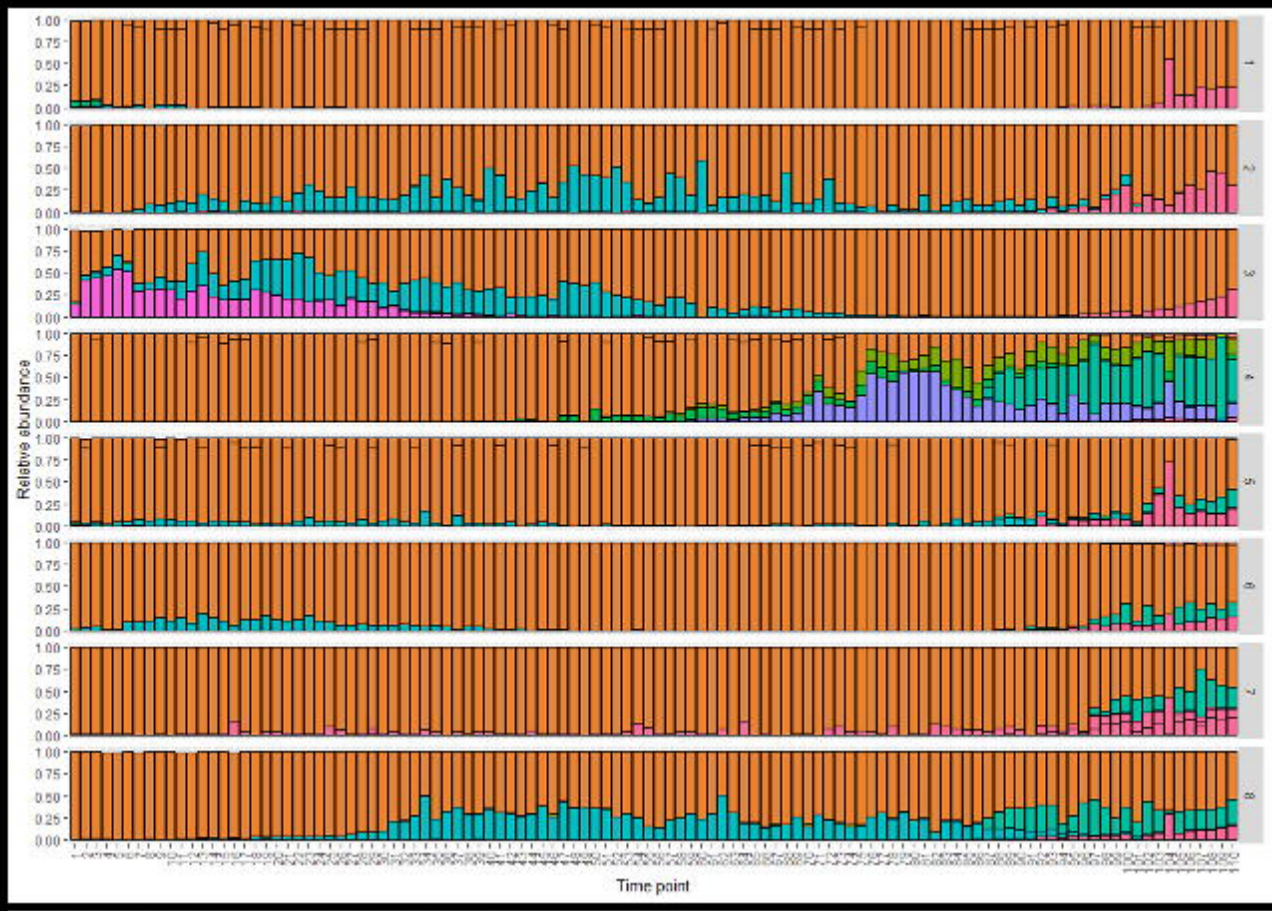
Response group
0
1
2
Phylum
Actinobacteriota
Bacteroidota
Firmicutes
Fusobacteriota
Proteobacteria
Component 1
Component 2
Subject index
Feature index
Time index
Fusobacterium canifelinum/nucleatum
Porphyromonas pasteri
Lachnoanaerobaculum cf.
Capnocytophaga leadbetteri
Aggregatibacter segnis
Gardiobacterium hominis
Streptococcus cristatus/sanguinis
Streptococcus cristatus/sinensis
Granulicatella elegans
Veillonella parvula/rogosae/tobetsuensis
Veillonella massiliensis
Leptotrichia hongkongensis
Fusobacterium massiliense
Actinomyces naeslundii/oris/viscosus/cellonella dispar/parvula
Kingella oralis
Prevotella nanceiensis
Rothia dentocariosa
Veillonella atypica/dispar
Feature mode, component 2 (transformed)
Feature mode, component 1 (transformed)
Cluster
1
2
3
Not clustered
Cluster 1
Cluster 2
Cluster 3
Sum of ASVs per group (mean +/- SEM)
Time point [days]
Response group
Low responders
High responders

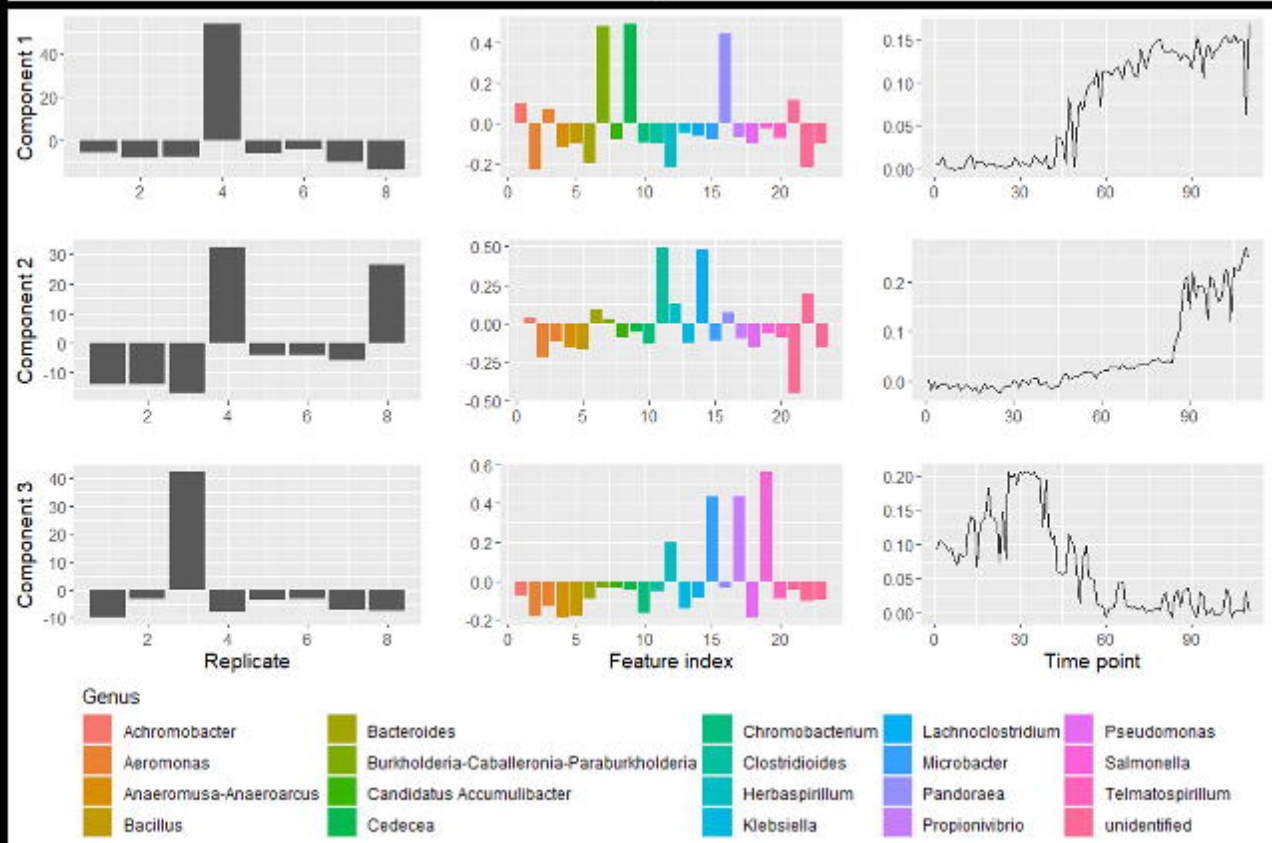

Time point 1
Time point K
Time point 1
X
E
X
E

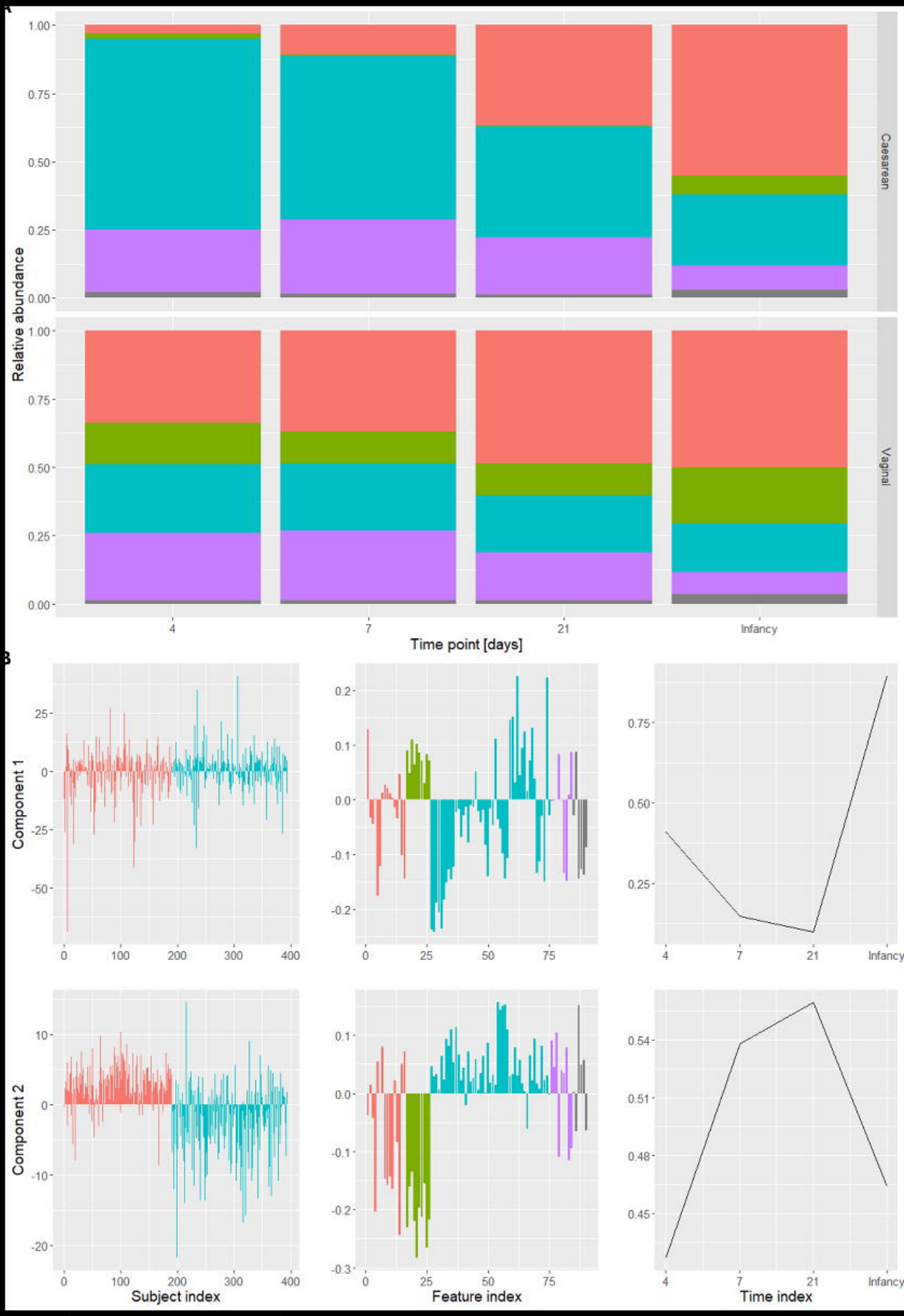

Cases
Controls
Relative abundance
Time point
Subject group
Case
Control
Phylum
Actinobacteria
Bacteroidetes
Proteobacteria


PCoA Component 2 (5%)
PCoA Component 1 (95%)
Phenotype
Case
Control
Time point
1
2
3
4
5
6
7
8
9
10


Component 1 loading
Component 2 loading
Subject index
Feature index
Time point


Relative abundance
Time point


Component 1
Component 2
Component 3
Replicate
Feature index
Time point
Genus
Achromobacter
Aeromonas
Anaeromusa-Anaeroarcus
Bacillus
Bacteroides
Burkholderia-Caballeronia-Paraburkholderia
Candidatus Accumulibacter
Cedecea
Chromobacterium
Clostridioides
Herbaspirillum
Klebsiella
Lachnoclostridium
Microbacter
Pandoraea
Propionivibrio
Pseudomonas
Salmonella
Telmatospirillum
unidentified


A
Relative abundance
Caesarean
Vaginal
Time point [days]
4
7
21
Infancy
B
Component 1
Component 2
Subject index
Feature index
Time index
Delivery mode
Caesarean
Vaginal
Phylum
Actinobacteria
Bacteroidetes
Firmicutes
Proteobacteria
NA