

<https://doi.org/10.1038/s41522-024-00565-x>

Multi-way modelling of oral microbial dynamics and host-microbiome interactions during induced gingivitis



G. R. van der Ploeg¹, B. W. Brandt², B. J. F. Keijser³, M. H. van der Veen^{2,4}, C. M. C. Volgenant^{2,5}, E. Zaura², A. K. Smilde¹, J. A. Westerhuis¹ & A. Heintz-Buschart¹✉

Gingivitis—the inflammation of the gums—is a reversible stage of periodontal disease. It is caused by dental plaque formation due to poor oral hygiene. However, gingivitis susceptibility involves a complex set of interactions between the oral microbiome, oral metabolome and the host. In this study, we investigated the dynamics of the oral microbiome and its interactions with the salivary metabolome during experimental gingivitis in a cohort of 41 systemically healthy participants. We use Parallel Factor Analysis (PARAFAC), which is a multi-way generalization of Principal Component Analysis (PCA) that can model the variability in the response due to subjects, variables and time. Using the modelled responses, we identified microbial subcommunities with similar dynamics that connect to the magnitude of the gingivitis response. By performing high level integration of the predicted metabolic functions of the microbiome and salivary metabolome, we identified pathways of interest that describe the changing proportions of Gram-positive and Gram-negative microbiota, variation in anaerobic bacteria, biofilm formation and virulence.

Microbial oral diseases, such as dental caries and gum inflammation, are a global public health burden^{1–3}. Together, they are the most prevalent health condition world-wide, impacting quality of life and imposing a significant economic burden with 5–10% of public health expenditure in most industrialized countries^{3,4}. Prevention of oral diseases is therefore of critical importance.

Gingivitis is a reversible stage of periodontal disease marked by inflamed, bleeding gums^{5–7}. It is caused by dental plaque accumulation due to poor oral hygiene^{8,9} and is present in a substantial part of the global population^{3,10}. The pathogenesis of gingivitis involves a complex set of interactions between the oral microbiome, the oral metabolome and the host^{5,11}. These are in turn dependent on many factors^{11,12}, including host diet¹³, life-style¹⁴, and genetics¹⁵. Therefore, the oral microbiome alone does not determine whether and to what extent gingivitis develops, as can be seen by the high degree of individual variation in gingivitis susceptibility seen in experimental gingivitis models after refraining from toothbrushing^{5,6,11,16–18}.

While there is no monocausal microbial agent in gingivitis, the gingival crevice microbiota underneath the gums shifts to a dysbiotic state^{7,19,20}. The

depletion of Gram-positive bacteria, such as *Rothia dentocariosa*, and enrichment in Gram-negative bacteria such as *Porphyromonas gingivalis*, *Prevotella* spp. and *Selenomonas* spp. is frequently observed in case-control comparisons, despite this shift being less pronounced compared to periodontitis progression^{5,7,21,22}. Additionally, increasing subgingival anaerobiosis during gingivitis progression promotes the growth of pathobionts^{7,23,24}.

Being the medium through which the host and the microbiota interact, the salivary metabolome is often studied as a non-invasive diagnostic of oral disease^{25,26}. A well-known host-microbiome interaction is the relation between host sugar consumption and the production of acids by the dental plaque microbiota. This lowers the local pH and shifts the microbiota further towards dysbiosis^{27,28}. However, many other dynamic microbe-microbe and host-microbiome interactions that occur throughout gingivitis onset and progression are not well understood.

Here, we investigate the spatially resolved oral bacterial microbiome and its interactions with the salivary metabolome during experimental gingivitis to (1) identify microbial subcommunities with common dynamics and (2) pinpoint potential impacts on biochemical pathways. Reversible

¹Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, The Netherlands. ²Preventive Dentistry, Academic Centre for Dentistry, Vrije Universiteit Amsterdam and University of Amsterdam, Amsterdam, The Netherlands. ³Microbiology and Systems Biology, TNO Healthy Living and Work, Leiden, The Netherlands. ⁴Paediatric Dentistry, Academic Centre for Dentistry, Vrije Universiteit Amsterdam and University of Amsterdam, Amsterdam, The Netherlands. ⁵Cariology, Academic Centre for Dentistry, Vrije Universiteit Amsterdam and University of Amsterdam, Amsterdam, The Netherlands. ✉e-mail: a.u.s.heintzbuschart@uva.nl

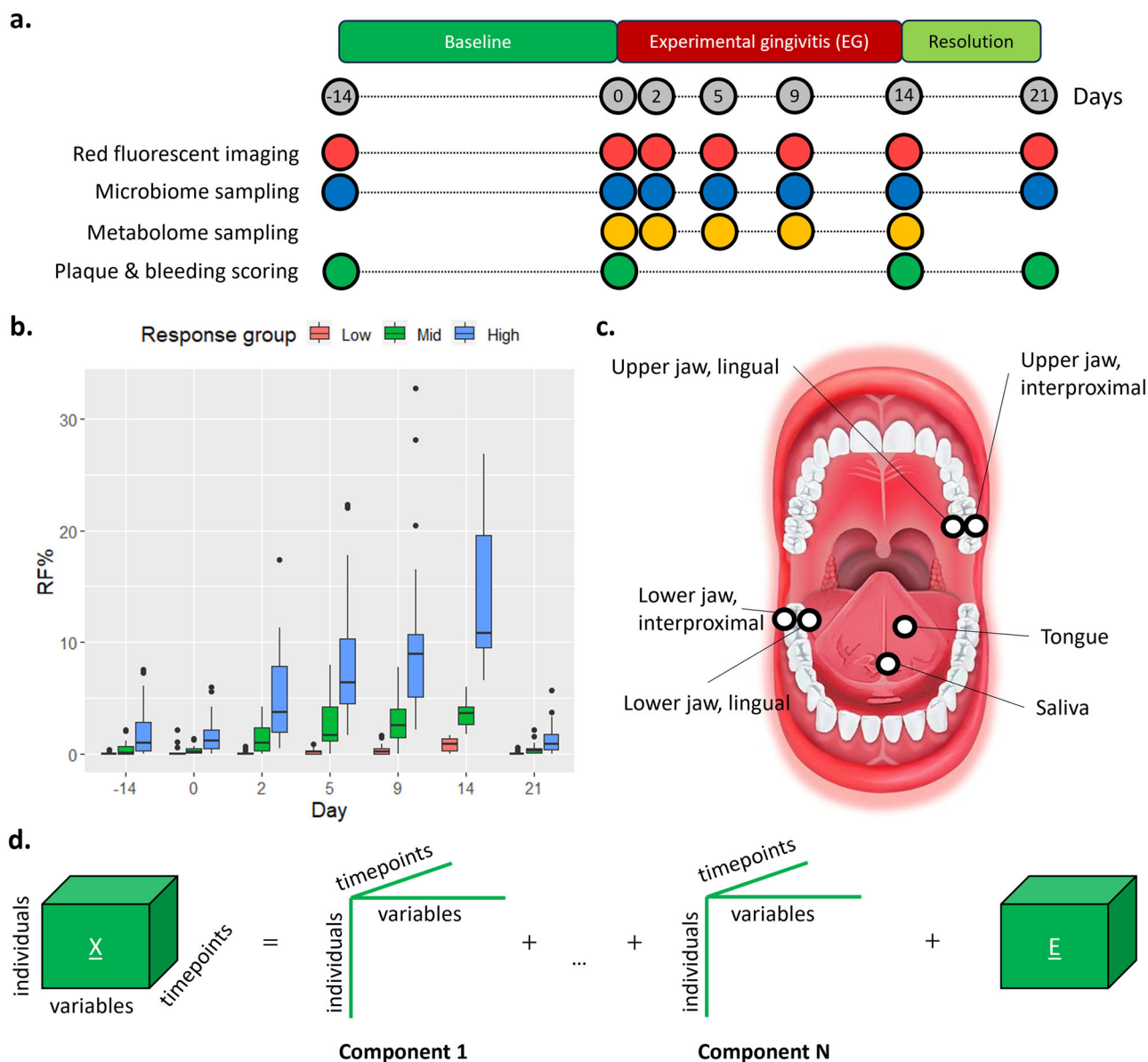


Fig. 1 | Overview of the study design and analysis. **a** Timeline of experimental gingivitis challenge with analyses. **b** Gingivitis responses: three groups of responders were identified based on red fluorescent dental plaque (RF%) at day 14: low (0–1.7%), mid (1.8–6.0%) and high (6.0%+). **c** Oral sampling sites (Image source: Ανώνυμος Βικιπαίδιστής, “The human mouth”, 2016. Accessed via https://commons.wikimedia.org/wiki/File:Human_mouth.jpg. CC BY-SA 3.0). **d** Data analysis

approach: Parallel factor analysis (PARAFAC), a multi-way generalization of Principal Component analysis, decomposes a multi-way data cube into components with three loading vectors each: one for the subjects, one for the variables (here: microbial relative abundances) and one for time. This approach allows us to identify the individual variation in the time-resolved response to the challenge and to describe commonly responding groups of microbiota.

gingivitis was induced in a cohort of 41 systemically healthy participants by omission of oral hygiene during a two-week gingivitis challenge^{18,29} (Fig. 1a). Previously reported plaque levels, gingival bleeding upon probing, and red fluorescent plaque quantification allowed grouping of the participants into identified low, mid and high responders based on the last day of the intervention¹⁸ (Fig. 1b). During a two-week baseline period at two time points (day -14 and day 0), a challenge period at four time points (day 2 to day 14), and one week resolution period at one time point (day 21), the oral microbiome was determined at six sites: tongue, saliva, supragingival plaque at the lower and upper lingual surfaces and supragingival plaque at the lower and upper interproximal surfaces (Fig. 1c). To identify the individual variation in the time-resolved response of the oral microbiome to the challenge and describe commonly responding (groups of) microbiota, we employ Parallel Factor Analysis (PARAFAC)^{30,31} for modelling the subjects, time, and microbial abundances at each site (Fig. 1d). The salivary metabolome

dynamics, assessed in unstimulated saliva during the challenge, is likewise modelled and integrated with microbiome models at the biochemical pathway level. These analyses should yield insights into microbial processes that underlie individual gingivitis responses.

Results

PARAFAC describes clinically relevant variation in oral microbiomes between subjects

To describe the microbial dynamics of each oral site for all subjects across time, we analysed 16S rRNA gene amplicon sequencing data for a total of 14,014 amplicon sequence variants (ASVs). After removal of sparse ASVs, we created separate PARAFAC models for each oral site, choosing the most appropriate number of PARAFAC components by inspecting the number of iterations needed to converge³², the core consistency diagnostic³³, the variation explained³², and the tucker

congruence coefficient per mode^{34,35} (Supplementary Data). This yielded two-component models for the upper jaw lingual plaque (19.7% explained variation) and upper jaw interproximal plaque (16.5%), tongue (31.5%) and saliva (17.5%) microbiomes (Fig. 2). Lower jaw lingual and interproximal plaque microbiomes were best represented by one-component models (explaining 11.4% and 7.2% of the variation, respectively). The model of the lower jaw interproximal plaque microbiome was not investigated further due to the limited amount of explained variation and due to the model not describing biologically meaningful information.

To interpret the models, we tested the correlation between the time-resolved subject loadings of each component with the measured parameters of gingivitis. This approach revealed correlation of components with one or more clinical parameter for the three plaque microbiomes (Table 1). Overall, the PARAFAC models describe the variation that exists between subjects, rather than the variation within a subject (one-sided Wilcoxon rank-sum test: $p = 0.020$, Supplementary Figs. 1 and 2). Furthermore, subject age was found to be significantly described by the first component of the model corresponding to the tongue microbiome samples ($p = 0.035$) and the second component of the model corresponding to the saliva microbiome samples ($p = 0.038$) (Supplementary Table 1). Gender was not significantly described by any model (Supplementary Table 1). We also performed Pearson correlation tests of the time-resolved subject loadings per component with richness and evenness. All microbiome models were found to contain at least one component that significantly described richness or evenness ($p \leq 0.05$; Table 1). In conclusion, the PARAFAC models represented clinically and ecologically relevant parameters.

PARAFAC describes microbial subcommunities with common gingivitis responses

Building on the interpretability of the subject modes (Table 1), we identified microbial subcommunities with common responses per sampling site by selecting and then clustering well-modelled ASVs based on their fitted response (Fig. 2, for cluster membership see Table 2). Per sampling site, at least one subcommunity containing mainly pathobionts or pro-inflammatory genera such as *Actinomyces*, *Campylobacter*, *Capnocytophaga*, *Fusobacterium*, *Leptotrichia* and *Porphyromonas*^{5,21,22,36,37} and one subcommunity containing mainly commensal genera such as *Kingella oralis*¹⁹, *Streptococcus* spp.^{38,39} and *Veillonella* spp.⁴⁰ were identified.

The sum of the relative abundances of the ASVs in each cluster was determined to test the difference between the microbiomes of individuals in the low and high response groups at every time point per oral site (Benjamini-Hochberg corrected permutation test of mean difference: $p \leq 0.05$; Supplementary Table 2). This approach revealed significant differences in the relative abundance of most plaque-associated ASV clusters in high responders compared to the low responders at baseline (day -14; Fig. 3, Supplementary Fig. 3). As such, the PARAFAC models of plaque microbiomes describe microbial subcommunities with common dynamics that connect to the magnitude of the gingivitis response.

We further investigated the robustness of the ecosystem during gingivitis onset and progression by comparing the sum of the relative abundances per cluster at the start and the end of the intervention (Supplementary Table 3). We found significant differences between the relative abundances of some ASV clusters at the start and end of the gingivitis challenge in lower jaw lingual (ASV cluster 1) and upper jaw interproximal (ASV cluster 4) plaque microbiomes in all response groups (Benjamini-Hochberg corrected two-sided Wilcoxon rank sum test: $p \leq 7.2e-13$ and $p \leq 2.6e-6$, respectively). For a further three ASV clusters (upper jaw lingual ASV cluster 1 and upper jaw interproximal clusters 1 and 3), we found the sum of relative abundances to be significantly different ($p = 0.002$, $p = 0.0025$, $p = 0.0039$, respectively) in high responders, but not in low responders. These results suggest that ecosystem stability is different between high and low responders in several oral niches.

Integration of the microbiome and metabolome PARAFAC models identifies host-microbiome interactions at the pathway level

With the established connection between the clinical parameters of gingivitis and the variation in the plaque microbiomes as modelled by PARAFAC, we investigated host-microbiome interactions at the biochemical pathway level. We obtained functional predictions of the microbiome per oral site using Tax4Fun2^{41,42}. We then created separate PARAFAC models of the salivary metabolomics data and of the functional prediction of the microbiomes at each oral site. Despite correcting for correlation between most salivary metabolite levels, likely due to variable water content of the saliva (see Methods), the subject loadings of both components of the model corresponding to the salivary metabolome were found to correlate with the correcting factor ($p = 2.0e-14$ and $p = 2.0e-5$, respectively). Regardless, we expected the salivary metabolomics model to have sufficient freedom to also describe variation related to the gingivitis response.

The well-modelled molecular functions (KEGG orthologous groups belonging to pathways that produce or consume the salivary metabolites) and well-modelled salivary metabolites were combined in a ranked list. KEGG-pathway enrichment was identified using SetRank, which is a functional enrichment algorithm that corrects for multiple pathway membership⁴³. This revealed that the microbiome and metabolome responses to the gingivitis challenge reflected differences in carbon and energy metabolism, as well as cell wall structure (Fig. 4, Supplementary Fig. 5, Supplementary Table 4). For example, the tongue microbiome and salivary metabolome were involved in a joint response in the biosynthesis of cofactors and carbon metabolism ($p = 0.0045$ and $p = 0.0052$, respectively). The quorum sensing pathway was significantly enriched in all plaque microbiomes ($p \leq 0.05$). The pathway enrichment results therefore reflect the microbiome responses and highlight potential metabolic interactions with the human host.

Discussion

We have shown that the unsupervised modelling approach of PARAFAC mainly described variation between subjects, as expected from an algorithm describing the maximum amount of variation that is present in the data. The modelled variation could be attributed to clinically relevant parameters. In addition, we have observed that the model corresponding to the salivary metabolomics samples partly describes the salivary dilution despite correcting for it using Probabilistic Quotient Normalization⁴⁴. This might be an artifact due to an incorrect assumption that correcting for water content of saliva can be done in the same way as for urine metabolomics samples. Future work on this topic should focus on the appropriateness of using Probabilistic Quotient Normalization prior to performing a decomposition analysis. Other multi-way approaches exist to describe a specific type of variation^{30–32}. For example, in N-way partial least squares (NPLS) the model is constrained to regress on an outcome variable⁴⁵. Additionally, the PARAFAC models of the microbiome and metabolome could be created in a linked fashion by keeping one of the modes equal to each other, known as coupled matrix and tensor factorization^{46–48}. Each of these approaches might yield new insights as the models explicitly look at variation that is common between the datasets. Further research is needed to compare such methods to PARAFAC in the context of multi-omics data.

We observed that the microbiome PARAFAC models partly describe richness and evenness. This result gives further evidence of a relationship between oral health and microbial diversity which has been described before^{49–51}. While the methodology used in this study targeted only bacteria, while other members of the oral microbiome—such as viruses and fungi—might also be relevant in gingivitis onset and progression^{52,53}. We identified several bacterial subcommunities within the oral sites that responded similarly to the intervention (Table 2). For example, a group of ASVs was significantly more abundant in individuals in the high responder group compared to low responders and represented species of known pathogenic or pro-inflammatory genera such as *Actinomyces*, *Campylobacter*, *Capnocytophaga*, *Fusobacterium*, *Leptotrichia*, and *Porphyromonas*^{5,21,22,36,37}.

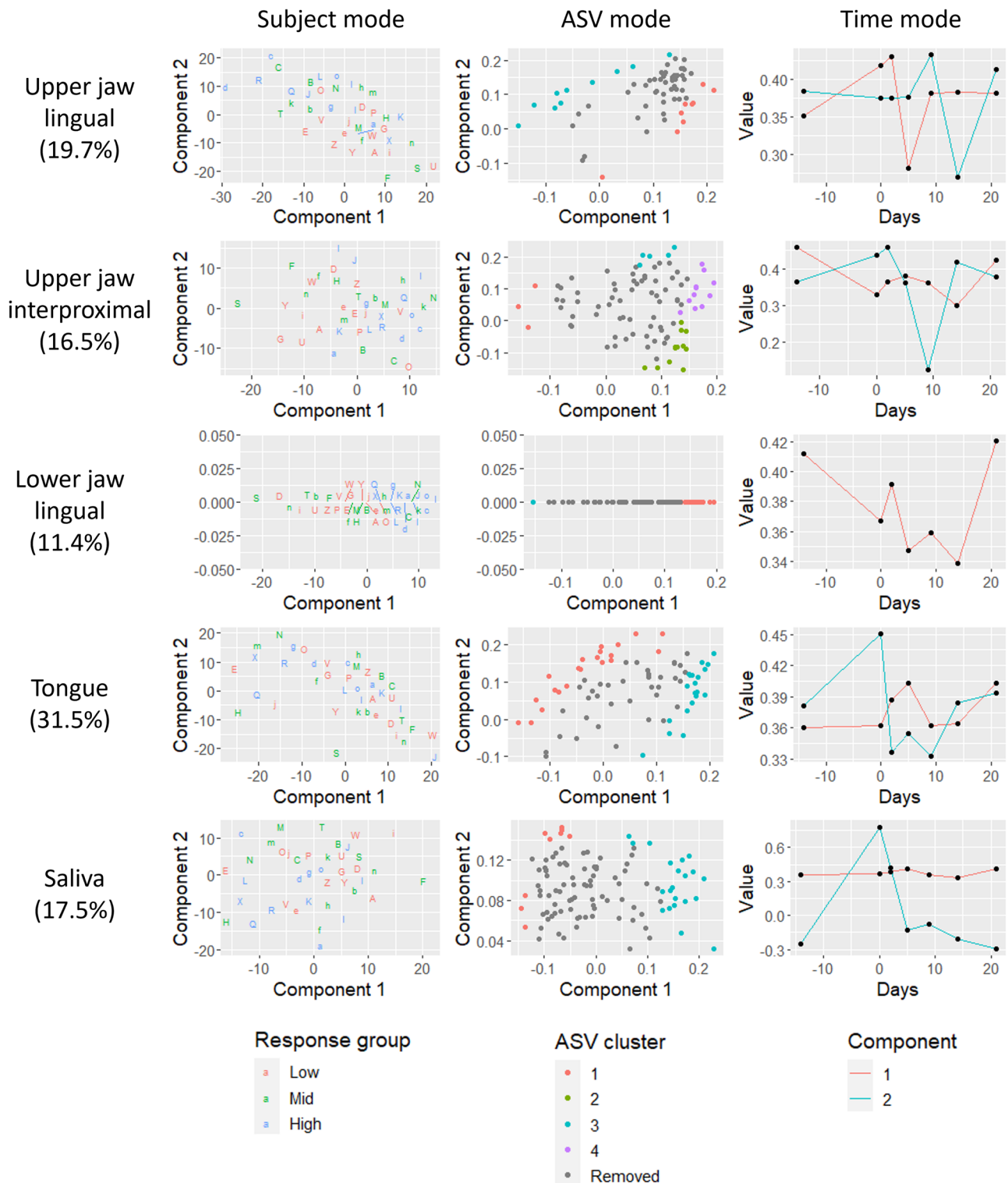


Fig. 2 | Overview of the PARAFAC models per sample type. In the rows, the PARAFAC models per sample type are described: upper jaw lingual plaque, upper jaw interproximal plaque, lower jaw lingual plaque, tongue and saliva. Variance explained per model are shown in the sample type labels. In the left column the loadings of the first and second component in the subject mode are plotted against each other, with every subject having a unique character to identify them across the sample types. In the middle column the loadings of the first and second component in the ASV mode are plotted against each other. We identified microbial

subcommunities (ASV clusters) with common responses per sampling site by selecting and then clustering well-modelled ASVs based on their fitted response (Methods, Table 2). ASVs are color-coded by cluster number (or shown in grey if removed prior to clustering). In the right column, the loadings per time point are shown per component. The model corresponding to the lower jaw lingual plaque only has one component. The model of the lower jaw interproximal microbiome was not investigated further due to the limited amount of explained variation and due to the model not describing biologically meaningful information.

Table 1 | Overview of the correlation test results between the time-resolved subject loadings of the PARAFAC models and the clinical parameters of gingivitis and microbiome diversity metrics

Sample type	Component	Plaque%	Bleeding%	RF%	Richness	Evenness
Upper jaw, lingual	1	0.024*	0.49	8.5e−6***	0.015*	0.032*
	2	0.002**	0.26	4.2e−7***	0.78	1.1e−6***
Upper jaw, interproximal	1	0.0015**	0.032*	8.1e−4***	4.4e−9***	7.0e−9***
	2	0.53	0.15	0.95	0.61	0.94
Lower jaw, lingual	1	4.9e−4***	0.14	1.2e−8***	1.6e−17***	1.5e−12***
Tongue	1	0.78	0.48	0.30	0.024*	0.057
	2	0.85	0.057	0.15	5.3e−11***	0.028*
Saliva	1	0.66	0.20	0.11	0.61	0.78
	2	0.78	0.81	0.78	7.3e−14***	8.5e−6***

Red fluorescence was assessed at every time point. Plaque and bleeding scores were only assessed at day -14, 0, 14 and 21 of the study. Plaque%: the percentage of sites in the mouth that were covered in plaque. Bleeding%: the percentage of sites in the mouth that bled upon probing. RF%: the percentage of sites that fluoresced red. Richness: the number of nonzero counts in a sample. Evenness: Shannon diversity divided by log(richness). Benjamini-Hochberg corrected p -values of Pearson correlation tests between the subject loadings of one component and the clinical parameters are reported (* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$).

Table 2 | Overview of all identified species per ASV cluster

Sample type	Cluster	Number of members	Identified species
Upper jaw, lingual	1	9	<i>Actinomyces naeslundii/oris/viscosus</i> , <i>Kingella oralis</i> , <i>Rothia dentocariosa</i> , <i>Streptococcus cristatis/sinensis</i> , <i>Veillonella dispar/parvula</i> , <i>Veillonella atypica/dispar</i> , <i>Veillonella parvula/rogosae/tobetsuensis</i>
	2	10	<i>Fusobacterium canifelinus/nucleatum</i> , <i>Fusobacterium massiliense</i> , <i>Granulicetella elegans</i> , <i>Prevotella nanceiensis</i> , <i>Veillonella massiliensis</i>
Upper jaw, interproximal	1	3	<i>Rothia dentocariosa</i> , <i>Veillonella atypica/dispar</i>
	2	12	<i>Abiotrophia defectiva</i> , <i>Aggregatibacter aphrophilus/kilianii</i> , <i>Capnocytophaga sputigena</i> , <i>Lachnoanaerobaculum cf.</i> , <i>Lautropia mirabilis</i> , <i>Rothia aerea/dentocariosa</i>
	3	5	<i>Corynebacterium matruchotii</i> , <i>Leptotrichia hofstadii</i> , <i>Leptotrichia shahii/wadei</i>
	4	11	<i>Aggregatibacter segnis</i> , <i>Capnocytophaga granulosa</i> , <i>Capnocytophaga leadberri</i> , <i>Cardiobacterium hominis</i> , <i>Fusobacterium canifelinus/nucleatum</i> , <i>Fusobacterium nucleatum</i> , <i>Leptotrichia buccalis</i> , <i>Porphyromonas pasteri</i>
Lower jaw, lingual	1	21	<i>Actinomyces georgiae/pacaensis</i> , <i>Aggregatibacter segnis</i> , <i>Campylobacter concisus</i> , <i>Capnocytophaga granulosa</i> , <i>Capnocytophaga leadbetteri</i> , <i>Cardiobacterium hominis</i> , <i>Corynebacterium matruchotii</i> , <i>Fusobacterium canifelinus/nucleatum</i> , <i>Fusobacterium nucleatum</i> , <i>Fusobacterium hwasooki/nucleatum/periodonticum</i> , <i>Lachnoanaerobaculum cf.</i> , <i>Leptotrichia hofstadii</i> , <i>Porphyromonas pasteri</i>
	2	1	<i>Streptococcus salivarius</i> (putative)
Tongue	1	21	<i>Actinomyces lingnae/marseillensis/pacaensis</i> , <i>Capnocytophaga leadbetteri</i> , <i>Fusobacterium periodonticum</i> , <i>Granulicatella adiacens/para-adiacens</i> , <i>Lachnoanaerobaculum cf.</i> , <i>Oribacterium parvum</i> , <i>Oribacterium sinus</i> , <i>Peptostreptococcus stomatis</i> , <i>Porphyromonas pasteri</i> , <i>Prevotella nanceiensis</i> , <i>Veillonella parvula/rogosae/tobetsuensis</i>
	2	20	<i>Actinomyces graevenitzi</i> , <i>Alloprevotella rava</i> , <i>Atopobium parvulum</i> , <i>Lachnoanaerobaculum orale/saburreum</i> , <i>Prevotella histicola</i> , <i>Prevotella jejuni/melaninogenica</i> , <i>Prevotella salivae</i> , <i>Prevotella veroralis</i> , <i>Rothia mucilaginosa</i> , <i>Stomatobaculum longum</i>
Saliva	1	9	<i>Granulicatella adiacens/para-adiacens</i> , <i>Porphyromonas pasteri</i> , <i>Rothia aerea/dentocariosa</i>
	2	21	<i>Actinomyces graevenitzi</i> , <i>Atopobium parvulum</i> , <i>Campylobacter concisus</i> , <i>Lachnoanaerobaculum orale/saburreum</i> , <i>Prevotella histicola</i> , <i>Prevotella jejuni/melaninogenica</i> , <i>Prevotella salivae</i> , <i>Rothia mucilaginosa</i> , <i>Stomatobaculum longum</i> , <i>Veillonella atypica/dispar</i>

ASVs within the same cluster show a similar response over time to the gingivitis intervention. Clustered ASVs are only reported here if two criteria were met: (1) the taxonomic classification was resolved to species level by DADA2 using the SILVA (v138) and HOMD (v15.22) databases and (2) the annotation agreed with a separate best hit annotation run using the HOMD (v15.22) database. Cases where the classification was resolved to species level by the DADA2 run but only to the genus level by the separate HOMD run are also reported.

Clusters of ASVs that were significantly more abundant in low responders compared to high responders contained known commensal species related to oral health, such as *Kingella oralis*¹⁹, *Streptococcus* spp.^{38,39}, and *Veillonella* spp.⁴⁰. It has to be noted that sequencing based only on the V4 region of the 16S rRNA gene may not provide sufficient resolution to correctly assign pathogenicity or pro-inflammatory traits to all ASVs. Similarly, some ASV have been assigned to bacterial taxa that are not commonly observed in the oral cavity. This may indicate that these bacteria were present in the sample due to problematic hygiene in the host or due to classification errors. Finally, we observed a difference in ecosystem stability between high and low

responders in some parts of the oral cavity. This suggests that in some subjects, the ecosystem can withstand the ecological pressure of plaque accumulation and has some mechanisms that prevents healthy microbiota from decreasing in abundance. Further research is needed to identify these mechanisms and the microbiota involved.

By integrating the PARAFAC models of the salivary metabolomics data and of the predicted metabolic functions from the microbiome data, we identified enriched pathways that reflected the changing proportions of Gram-positive and -negative microbiota (cell membrane/wall component pathways, such as lipopolysaccharide and peptidoglycan biosynthesis).

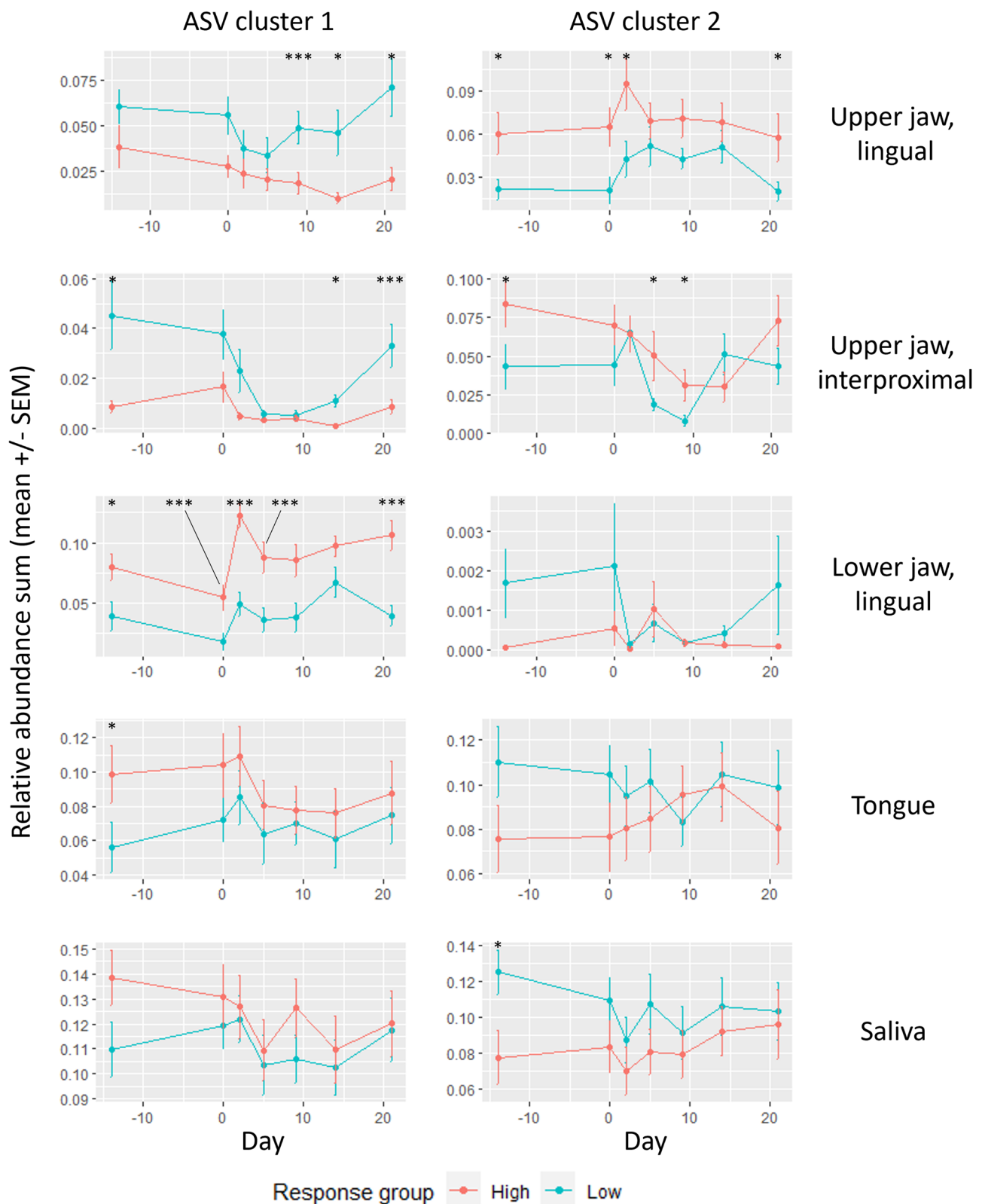


Fig. 3 | Overview of the sums of relative abundances per ASV cluster and sample type, separating subjects by response group. Error bars correspond to the standard error of the mean (SEM). The mean difference between the high and low response groups per time point was tested using a Benjamini-Hochberg corrected

permutation test of 999 iterations (* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$). Mid responders and ASV clusters 3 and 4 of the upper jaw interproximal plaque samples are not shown for visual clarity (Supplementary Figure 3 and 4, respectively). Please refer to Table 2 for the list of identified species per ASV cluster.

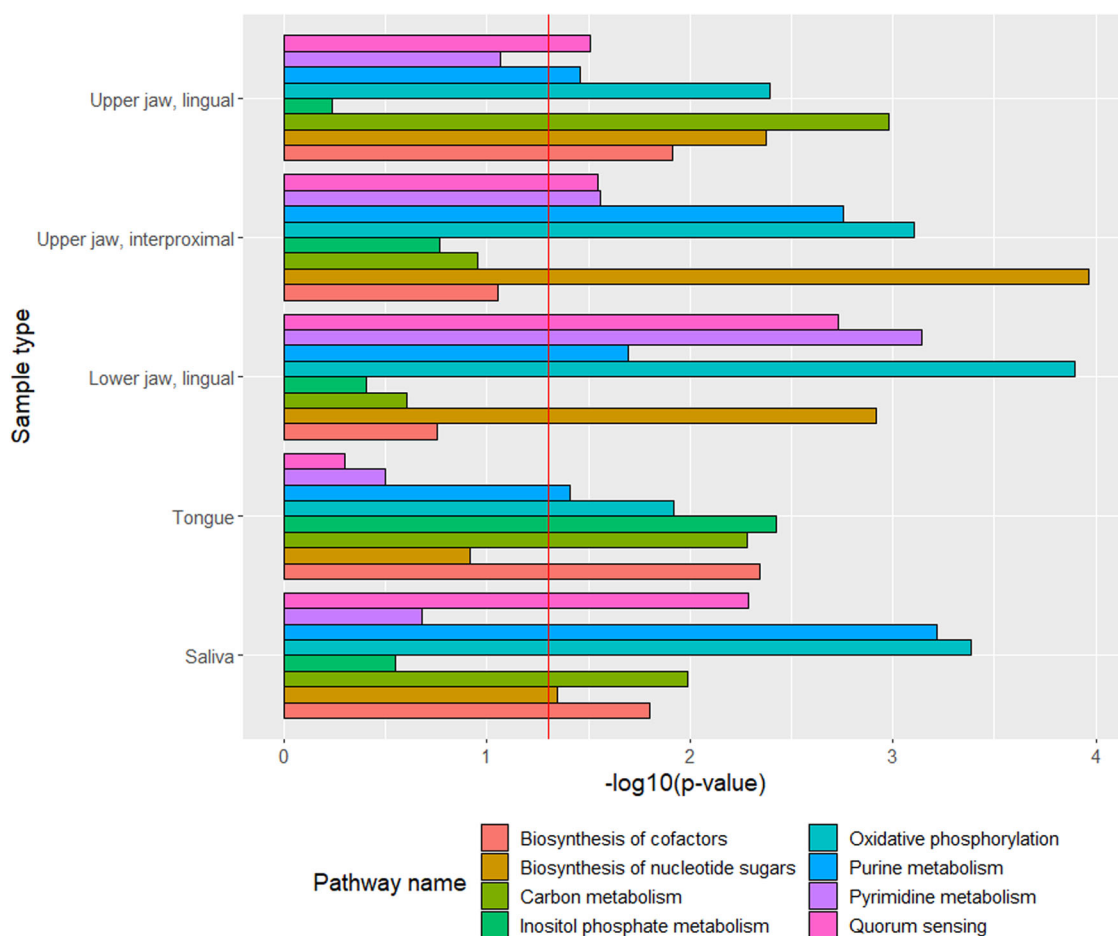


Fig. 4 | Overview of the pathway enrichment results per combination of oral site microbiome and the salivary metabolome. Tax4Fun2 was used to create functional predictions from ASV data. Well-modelled metabolites and microbiome molecular functions were integrated by mapping them to KEGG pathway level. SetRank was

then used to test for pathway enrichment and corrected for multiple-pathway membership. Pathways were filtered to be significant ($p < 0.01$) in at least one sample type and to contain at least two well-modelled KEGG molecular functions and two well-modelled metabolites. The red vertical line corresponds to $p = 0.05$.

Quorum sensing systems were found to be significantly enriched in all plaque sites likely due to the presence of various *Streptococcus* spp., that carry genes for oligopeptide binding protein quorum sensing system^{54,55}. These are known to regulate biofilm formation and virulence^{56–58}. The strong signal in carbon metabolism and oxidative phosphorylation were likely reflective of the variation in anaerobic bacteria, which accumulate during gingivitis^{59–61}. An interesting link between salivary metabolites and microbiome functional potential may have been observed in the nucleoside/purine/pyrimidine biosynthesis pathways: while urea was one of the best modelled salivary metabolites and is a potential biomarker for oral health^{61–64}, a role of salivary nucleosides in plaque formation or gingivitis is as yet unknown. While many of these functions can be identified as housekeeping activities, we protected ourselves from wrongfully identifying such generic activities as enriched by (1) using SetRank to correct for multiple pathway membership and (2) creating a custom database of pathway elements and removing very large pathways that often connect to such activities. Further research is needed to investigate the relevance of these enriched pathways in the context of dental plaque formation and gingivitis.

In conclusion, we show that PARAFAC modelling of longitudinal oral microbiome and salivary metabolomics data can be used to identify clinically relevant variation and similarly responding microbial subcommunities related to gingivitis. By performing high level integration of predicted microbiome functions and the salivary metabolome, we highlight biochemical pathways of oral biofilm formation and maturation with a likely role in gingivitis.

Methods

Data description

The gingivitis challenge study was carried out at the Academic Centre for Dentistry Amsterdam (ACTA) with 15 systemically healthy males and 26 systemically healthy females between the ages of 18 and 55. Recruitment details and exclusion criteria were previously described²⁹. Reversible gingivitis was induced by omission of oral hygiene during the two-week gingivitis challenge (day 0 to day 14), which was preceded by a two-week baseline period (day -14 to day 0) and followed by a one-week resolution phase (day 14 to day 21; Fig. 1a). Assessment of plaque formation, bleeding and the collection of oral samples for microbiome were performed throughout the baseline, challenge and resolution phases. Plaque and bleeding were assessed clinically in a half mouth randomized contralateral model⁶⁵. Oral samples for microbiome analysis were taken at six sites in the mouth: supragingival plaque at the lower and upper jaw lingual surfaces, supragingival plaque at the lower and upper jaw interproximal surfaces, tongue and saliva. The salivary metabolome was only sampled during the challenge phase.

Red fluorescence imaging

Acquisition of red fluorescent plaque (RF) photographs has been described previously^{18,66,67}. Briefly, fluorescence photographs were taken of the vestibular aspect of the anterior teeth (cuspid to cuspid, upper and lower jaw) in end-to-end position at every time point using a QLF-D camera (Inspektor Research Systems BV, Amsterdam, the Netherlands). The photographs were assessed planimetrically for the percentage red fluorescent protein (RFP) coverage using RFP analysis software (QA2 V1.25, Inspektor

Research Systems BV, Amsterdam, the Netherlands). Three response groups were determined using the RF⁶ values on day 14: low (0–1.7%), mid (1.8–6.0%) and high (6.0%+) (Fig. 1b)¹⁸.

DNA isolation and 16S rRNA gene amplicon sequencing

DNA was isolated from the samples using a previously described method⁶⁸. Samples were mixed with 300 µl lysis buffer (Agowa, Berlin, Germany), 500 µl phenol saturated with tris-HCl (pH 8.0) and 500 µl zirconium beads (0.1 mm; BioSpec Products, Bartlesville, OK, USA), and shaken in a bead beater for 3 min at 2800 oscillations/min. DNA released was purified using magnetic beads (Agowa, Berlin, Germany) and used for amplicon sequencing.

The V4 hypervariable region of the 16S rRNA gene was targeted using primers F515 (5'- GTG CCA GCM GCC GCG GTA A -3') and R806 (5'- GGA CTA CHV GGG TWT CTA AT -3'). The primers included Illumina adapters and a unique 8-nucleotide sample index sequence key⁶⁹. PCR was performed using the Phusion Hot Start II High Fidelity PCR Master Mix (Thermo Scientific, Waltham, MA, USA) with 100 pg template. The following amplification program was used: initial denaturation at 98 °C for 30 s; 30 cycles of 98 °C for 10 s, 55 °C for 30 s and 72 °C for 30 s; and final elongation at 72 °C for 5 min. The amplicon libraries were pooled in equimolar amounts and purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA). Amplicon quality and size were analysed on a Fragment Analyzer (Advanced Analytical Technologies Inc., Ankeny, IA, USA). Amplicon sequencing was performed on the Illumina MiSeq platform (Illumina Technologies, San Diego, CA, USA) using 2 × 200 cycle paired-end settings.

16S rRNA gene amplicon data pre-processing

The 16S rRNA gene sequencing data were pre-processed using DADA2⁷⁰ (version 1.14.0). All microbiome sampling locations except for saliva were pre-processed with default DADA2 settings (truncLen = c(100, 100), maxN = 0, maxEE = c(2,2), truncQ = 2, rm.phix = TRUE, compress = FALSE, matchIDs = TRUE, multithread = TRUE). The saliva samples were processed with a shorter trimming setting (truncLen = c(200, 215)) due to lower quality reads. Taxonomic classification of ASVs was performed using the SILVA (v138) and HOMD (v15.22) reference databases.

Metabolomics sampling and profiling

Saliva sample collection and metabolite profiling were described previously²⁹. Unstimulated saliva was collected at the intervention time points. Participants were instructed to allow saliva to accumulate on the floor of the mouth and to spit at 30 second intervals into a pre-weighted 30 ml polypropylene tube. The collection period was 5 minutes. Non-targeted metabolite profiling of the saliva samples for the gingivitis challenge time points was performed by Metabolon. Samples were processed as described in Metabolon's standard method⁷¹. Raw data were extracted, peak-identified and quality controlled using Metabolon's hardware and software. Compounds were identified by comparison with Metabolon's reference library⁷². Metabolites were annotated with KEGG compound IDs^{73,74}.

Assessment of clinical parameters of gingivitis

Plaque and bleeding scores were clinically assessed by two experts in a half mouth randomized contralateral model, as described previously¹⁸. Plaque was quantified using a modified Silness & Loe Plaque Index at six sites of the buccal and lingual aspects of all present teeth⁷⁵. Gingival bleeding was quantified using the bleeding on marginal probing index on six gingival areas on the buccal and lingual sides of all present teeth⁷⁶.

Statistical analysis—microbiome data processing

The ASV data and taxonomic information were processed using MATLAB (version 2022a). The data were separated by oral site. To limit the sparsity per dataset, ASVs were kept if the sparsity was <50% in any response group (Supplementary Fig. 6). All other ASVs were removed from the data. By

setting a response group-based sparsity-cutoff per ASV, we avoid removing biologically relevant ASVs that would occur in only one or two response groups. ASVs were also removed if they corresponded to chloroplast or mitochondrial sequences, as these were not relevant for the study. An overview of the number of ASVs before and after filtering per sample type is reported in Supplementary Table 5. Subsequently, a centered-log ratio transformation—using a pseudo-count of 1—was performed to correct for compositionality^{77,78}. The datasets were then converted to three-way arrays (Fig. 1d). Missing samples were kept in the data cube as a row of missing values (Supplementary Table 6). This is possible because PARAFAC interpolates the missing data automatically in its alternating least-squares algorithm and maximises the amount of information for the modelling procedure as the other samples of the subject do not have to be removed entirely. Subsequently, centering across the subject mode and scaling within the ASV mode—ignoring missing values—were performed to make the samples comparable at every timepoint and the ASVs comparable for all time points⁷⁹.

Statistical analysis—salivary metabolomics data processing

The salivary metabolomics data were processed using MATLAB (version 2022a). To limit the sparsity per dataset, metabolites were excluded if they contained more than 25% values below the detection limit (Supplementary Fig. 7). Additionally, xenobiotic compounds were manually assessed for occurrence across response groups and selected if they were prevalent in most subjects (Supplementary Data). After feature selection, 400 of the 499 metabolites remained (Supplementary Table 5). Values below the detection limit were imputed with a random value between 0 and the detection limit per metabolite to preserve their distribution. To correct for the dilution caused by the amount of water in the saliva samples, Probabilistic Quotient Normalization (PQN) was performed using the median value of each metabolite as an artificial reference sample⁴⁴. Next, the dataset was (natural) log transformed to stabilize the variance. The dataset was then converted to a three-way array. Subsequently, centering across the subject mode and scaling within the ASV mode—ignoring missing values—were performed to make the samples comparable at every timepoint and the ASVs comparable for all time points⁷⁹. All subjects were fully sampled, except for subject 3CN8CB for whom no metabolome data was available (Supplementary Table 6).

Statistical analysis—Functional profile prediction of microbiome data

A functional profile prediction based on the microbiome ASV count data was performed using the Tax4Fun2 package in R (Tax4Fun2 version 1.1.5, R version 4.0.3)^{41,42}. ASVs corresponding to Chloroplast or Mitochondria were removed. Samples were rarefied to 10 000 reads to make them comparable (Supplementary Figs. 8 and 9). Samples with fewer total reads were removed. This step removed 53 of 1692 samples. ASVs without counts after rarefying were removed. All steps together removed 6903 out of 14,014 ASVs. Tax4Fun2 was run with default settings against the included Ref99NR database⁴² and a custom database based on the HOMD genomes (v9.14, accessed 2020-11-09)⁸⁰. Molecular functions of the genomes were assigned using Tax4Fun2's DIAMOND wrapper⁸¹. HOMD-annotated 16S rRNA gene sequences were extracted from the genomes, cut to the 515F-806R fragment used in this study using cutadapt v1.18⁸². Exact duplication within genomes were removed, and 16S rRNA gene sequence fragments were combined with the functional assignments using Tax4Fun's generateUserDataByClustering function. The resulting functional prediction profiles—which contain values between 0 and 1 for each function—were used for subsequent analysis. The fraction of unused taxonomic units and the fraction of unused sequences are reported per sample (Supplementary Figs. 10 and 11).

Statistical analysis—processing of functional profile predictions

The functionally predicted microbiome data were processed in MATLAB (version 2022a). The data were separated by sample type. To make the data

comparable to the salivary metabolomics data, predicted KEGG orthologous groups (KOs) were removed if they did not belong to pathways that the metabolites mapped to. This was done by mapping the KOs and metabolite compound IDs to pathways using the KEGG API⁸³. To limit the sparsity per dataset, KOs were removed if the number of zeroes for all response groups was >50% (Supplementary Fig. 12). Additionally, KOs were removed if the sum-of-squares was lower than 0.025% of the total sum-of-squares of the dataset (Supplementary Fig. 13). This latter filter was to ensure that the modelling procedure focused on relevant variation in the data and made the total number of variables comparable to the metabolomics data, while retaining most information. An overview of the number of KOs remaining after feature selection is reported (Supplementary Table 5). Subsequently a centered-log ratio transformation of each dataset was performed to correct for compositionality^{77,78}. The dataset was then converted to a three-way array. Missing samples were kept in the data cube as a row of missing values (Supplementary Table 6). Subsequently, centering across the subject mode and scaling within the ASV mode - ignoring missing values—were performed to make the samples comparable at every timepoint and the ASVs comparable for all time points⁷⁹.

Statistical analysis—Parallel Factor Analysis (PARAFAC)

Details on the creation of PARAFAC models for microbiome⁸⁴ and metabolomics data³² have been described elsewhere. The PARAFAC implementation from the N-way toolbox (version 1.8.0.0, <https://nl.mathworks.com/matlabcentral/fileexchange/1088-the-n-way-toolbox>) in MATLAB (version 2022a) was used to create PARAFAC models for all datasets. Similar to PCA, the correct number of components of the PARAFAC model needed to be determined to create an optimal model^{33,85}. This was done by inspecting the number of iterations needed to converge³², the core consistency diagnostic (CORCONDIA)³³, the variation explained³², and the Tucker congruence coefficient per mode^{34,35}. Additionally, a jack-knife approach was used to determine the stability of the PARAFAC modelling procedure. These metrics were inspected per component for every generated model (Supplementary Data). All generated models are reported in Supplementary Figs. 14–26. An overview of the model statistics also reported (Supplementary Table 7). Congruence loadings, which describe the relationships between the original variables of the dataset and the latent variables of the corresponding model, were calculated for every component in every sample type⁸⁶. While the metrics above suggested a one-component PARAFAC model corresponding to the microbiome data at the lower jaw interproximal niche, the model described less than 10% of the variation in the data. Furthermore, manual inspection revealed that the model did not describe variation that we could explain biologically. Hence, we did not analyse this model further in subsequent steps.

A transformation of the subject loadings was required for comparison with the longitudinally measured clinical parameters of gingivitis. The ASV loading vectors were orthonormalized using the Gram-Schmidt orthonormalization procedure in the *pracma* R package^{87,88} (version 2.4.2). The (same) transformation matrix is then applied to the Kronecker product of the subject and the time loadings to obtain interpretable loadings of every subject-time combination⁸⁹. The correlation of the time-resolved subject loadings with the clinical parameters of gingivitis was then tested (Table 1 and Supplementary Table 1).

Statistical analysis—microbiome ASV cluster analysis

Loading plots of the ASV mode and the subject mode for each of the microbiome PARAFAC models were created (Fig. 2). ASVs were filtered out if they had a variation explained lower than the average of the model or a congruence loading lower than 0.4⁸⁶. ASVs were clustered based on their fitted abundances according to the PARAFAC model using the K-medoids algorithm from the cluster R package (version 2.1.4) with 50 random starts to be robust against outliers⁹⁰. The number of clusters was determined using the within-cluster sum of squares, silhouette width and gap statistic metrics as reported by the *factoextra* R package⁹¹ (version 1.0.7, Supplementary

Fig. 27). An overview of the taxonomic information per cluster can be found in Table 2. Bacteria in Table 2 were only reported if two criteria were met: (1) the taxonomic classification was resolved to species level by DADA2 using the SILVA (v138) and HOMD (v15.22) databases and (2) the annotation agreed with a separate best hit annotation run using the HOMD (v15.22) database. Cases where the classification was resolved to species level by the DADA2 run but only to the genus level by the separate HOMD run are also reported.

The response to the gingivitis intervention of the ASV clusters was shown using the relative abundance sum per ASV cluster derived from the original count data (Fig. 3 and Supplementary Fig. 3). The mean of the summed relative abundances and standard error of the mean were calculated per response group for a given sample type and time point. The mean difference between the high and low responders was tested using a permutation analysis where the response group membership of the subjects was permuted (Supplementary Table 2).

Statistical analysis—functional enrichment analysis

The PARAFAC models of the functionally profiled microbiome and salivary metabolomics data were combined to perform a pathway enrichment test per sample type. This was done by calculating the variance explained and congruence loadings per KO or metabolite in each model⁸⁶. If the PARAFAC model contained two components, the maximum of the two congruence loadings per feature was used. The variance explained and congruences were normalized by the maximum value per model prior to integration. The normalized variance explained and normalized congruence per feature were averaged to obtain a ranking of pathway elements (KOs and metabolites). This causes the best modelled KO and compound ID to be at the top of this list. The ranked pathway element list was cut off at 75% of its length to ensure a good separation between well-modelled and averagely-modelled pathway elements (Supplementary Fig. 28).

The SetRank R package (version 1.1.0) was used to obtain pathways of interest⁴³. SetRank is a gene set enrichment algorithm that obtains a high sensitivity by correcting for multiple pathway membership. A custom database of pathway elements was created by combining the KO and compound ID to pathway mappings obtained through the KEGG API (Supplementary Data). Pathways were removed from our database if they could not be performed by prokaryotes. Generic, large pathways with more than 750 elements were removed to ensure that the enrichment results were specific enough for interpretation (Supplementary Fig. 29). Subsequently the SetRank analysis was performed with standard settings per sample type. The multiple-pathway corrected *p*-values are reported (*p* ≤ 0.05, Fig. 4 and Supplementary Table 4).

Data availability

The ASV abundance data is available at <https://github.com/GRvanderPloeg/TIFN-multiway>. Raw sequencing data is available by request from the authors, in accordance with the informed consent signed by the study participants.

Code availability

The underlying code for this study is available on GitHub and can be accessed via <https://github.com/GRvanderPloeg/TIFN-multiway/releases/tag/v1.2>.

Received: 7 April 2024; Accepted: 12 September 2024;

Published online: 19 September 2024

References

- Petersen, P. E., Bourgeois, D., Ogawa, H., Estupinan-Day, S. & Ndiaye, C. The global burden of oral diseases and risks to oral health. *Bull. World Health Organ.* (2005).
- Peres, M. A. et al. Oral diseases: a global public health challenge. *Lancet* **394**, 249–260 (2019).

3. WHO. Global Oral Health Status Report: *Towards Universal Health Coverage for Oral Health by 2030*. <https://www.who.int/publications/i/item/9789240061484> (2022).
4. Widström, E., Eaton, K. & Vanobbergen, J. Oral healthcare systems in the Extended European Union, partim:[Oral Health care system in] Belgium. *Oral. Health Prev. Dent.* **2**, 155–157 (2004).
5. Huang, S. et al. Predictive modeling of gingivitis severity and susceptibility via oral microbiota. *ISME J.* **8**, 1768–1780 (2014).
6. Chapple, I. L. C. et al. Primary prevention of periodontitis: managing gingivitis. *J. Clin. Periodontol.* **42**, S71–S76 (2015).
7. Curtis, M. A., Diaz, P. I. & Van Dyke, T. E. The role of the microbiota in periodontal disease. *Periodontol 2000* **83**, 14–25 (2020).
8. Löe, H., Theilade, E. & Jensen, S. B. Experimental Gingivitis in Man. *J. Periodontol.* **36**, 177–187 (1965).
9. Murakami, S., Mealey, B. L., Mariotti, A. & Chapple, I. L. C. Dental plaque-induced gingival conditions. *J. Clin. Periodontol.* **45**, S17–S27 (2018).
10. Han, L. et al. Hygiene practices among young adolescents aged 12–15 years in low- and middle-income countries: a population-based study. *J. Glob. Health* **10**, 020436 (2020).
11. Chapple, I. L. C. et al. Periodontal health and gingival diseases and conditions on an intact and a reduced periodontium: Consensus report of workgroup 1 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *J. Periodontol.* **89**, S74–S84 (2018).
12. Kilian, M. et al. The oral microbiome – an update for oral healthcare professionals. *Br. Dent. J.* **221**, 657–666 (2016).
13. Van Der Velden, U., Kuzmanova, D. & Chapple, I. L. C. Micronutritional approaches to periodontal therapy. *J. Clin. Periodontol.* **38**, 142–158 (2011).
14. Bergström, J. & Preber, H. The influence of cigarette smoking on the development of experimental gingivitis. *J. Periodontol. Res.* **21**, 668–676 (1986).
15. Nibali, L., Di Iorio, A., Tu, Y. & Vieira, A. R. Host genetics role in the pathogenesis of periodontal disease and caries. *J. Clin. Periodontol.* **44**, (2017).
16. Axelsson, P., Lindhe, J. & Nyström, B. On the prevention of caries and periodontal disease. *J. Clin. Periodontol.* **18**, 182–189 (1991).
17. Guk, H.-J., Lee, E.-S., Jung, U.-W. & Kim, B.-I. Red fluorescence of Interdental plaque for screening of gingival health. *Photodiagnosis Photodyn. Ther.* **29**, 101636 (2020).
18. van der Veen, M. H., Volgenant, C. M. C., Keijser, B., ten Cate, J., Bob, M. & Crielaard, W. Dynamics of red fluorescent dental plaque during experimental gingivitis—A cohort study. *J. Dent.* **48**, 71–76 (2016).
19. Abusleme, L., Hoare, A., Hong, B. & Diaz, P. I. Microbial signatures of health, gingivitis, and periodontitis. *Periodontol 2000* **86**, 57–78 (2021).
20. Diaz, P. I., Hoare, A. & Hong, B.-Y. Subgingival Microbiome Shifts and Community Dynamics in Periodontal Diseases. *J. Calif. Dent. Assoc.* **44**, 421–435 (2016).
21. Kistler, J. O., Booth, V., Bradshaw, D. J. & Wade, W. G. Bacterial Community Development in Experimental Gingivitis. *PLoS ONE* **8**, e71227 (2013).
22. Schincaglia, G. P. et al. Clinical, Immune, and Microbiome Traits of Gingivitis and Peri-implant Mucositis. *J. Dent. Res.* **96**, 47–55 (2017).
23. Diaz, P. I., Zilm, P. S. & Rogers, A. H. *Fusobacterium nucleatum* supports the growth of *Porphyromonas gingivalis* in oxygenated and carbon-dioxide-depleted environments. *Microbiology* **148**, 467–472 (2002).
24. Ter Steeg, P. F., Van Der Hoeven, J. S., De Jong, M. H., Van Munster, P. J. J. & Jansen, M. J. H. Modelling the Gingival Pocket by Enrichment of Subgingival Microflora in Human Serum in Chemostats. *Microb. Ecol. Health Dis.* **1**, 73–84 (1988).
25. Dawes, C. & Wong, D. T. W. Role of Saliva and Salivary Diagnostics in the Advancement of Oral Health. *J. Dent. Res.* **98**, 133–141 (2019).
26. Proctor, G. B. The physiology of salivary secretion. *Periodontol 2000* **70**, 11–25 (2016).
27. König, K. G. & Navia, J. M. Nutritional role of sugars in oral health. *Am. J. Clin. Nutr.* **62**, 275S–282S (1995).
28. Marsh, P. D. Sugar, fluoride, pH and microbial homeostasis in dental plaque. *Proc. Finn. Dent. Soc. Suom. Hammaslaakariseuran Toim.* **87**, 515–525 (1991).
29. Prodan, A. et al. Effect of experimental gingivitis induction and erythritol on the salivary metabolome and functional biochemistry of systemically healthy young adults. *Metabolomics* **12**, 147 (2016).
30. Carroll, J. D. & Chang, J.-J. Analysis of individual differences in multidimensional scaling via an n-way generalization of “Eckart-Young” decomposition. *Psychometrika* **35**, 283–319 (1970).
31. Harshman, R. A. Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multimodal factor analysis. UCLA Working Papers in Phonetics 16, 1–84 (1970).
32. Bro, R. PARAFAC. Tutorial and applications. *Chemom. Intell. Lab. Syst.* **38**, 49–171 (1997).
33. Bro, R. & Kiers, H. A new Efficient Method for Determining the Number of Components in PARAFAC Models. *J. Chemom.* **17**, 274–286 (2003).
34. Lorenzo-Seva, U. & ten Berge, J. M. F. Tucker’s congruence coefficient as a meaningful index of factor similarity. *Methodol. Eur. J. Res. Methods Behav. Soc. Sci.* **2**, 57–64 (2006).
35. Tucker, L. R. *A Method for Synthesis of Factor Analysis Studies*. 984 (Educational Testing Service Princeton, NJ, 1951).
36. Kirst, M. E. et al. Dysbiosis and Alterations in Predicted Functions of the Subgingival Microbiome in Chronic Periodontitis. *Appl. Environ. Microbiol.* **81**, 783–793 (2015).
37. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
38. Caporaso, J. G. et al. Moving pictures of the human microbiome. *Genome Biol.* **12**, R50 (2011).
39. Zaura, E., Keijser, B. J., Huse, S. M. & Crielaard, W. Defining the healthy” core microbiome” of oral microbial communities. *BMC Microbiol.* **9**, (2009).
40. Moore, W. E. C. & Moore, L. V. H. The bacteria of periodontal diseases. *Periodontol 2000* **5**, 66–77 (1994).
41. Aßhauer, K. P., Wemheuer, B., Daniel, R. & Meinicke, P. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* **31**, 2882–2884 (2015).
42. Wemheuer, F. et al. Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. *Environ. Microbiome* **15**, 11 (2020).
43. Simillion, C., Liechti, R., Lischer, H. E. L., Ioannidis, V. & Bruggmann, R. Avoiding the pitfalls of gene set enrichment analysis with SetRank. *BMC Bioinforma.* **18**, 151 (2017).
44. Dieterle, F., Ross, A., Schlotterbeck, G. & Senn, H. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. *Anal. Chem.* **78**, 4281–4290 (2006).
45. Bro, R. Multiway calibration. Multilinear PLS. *J. Chemom.* **10**, 47–61 (1996).
46. Acar, E. et al. Structure-revealing data fusion. *BMC Bioinforma.* **15**, 239 (2014).
47. Acar, E., Bro, R. & Smilde, A. Data Fusion in Metabolomics Using Coupled Matrix and Tensor Factorizations. *Proc. IEEE* **103**, 1602 (2015).
48. Singh, A. P. & Gordon, G. J. Relational learning via collective matrix factorization. in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* 650–658 (Association for Computing Machinery, New York, NY, USA, 2008). <https://doi.org/10.1145/1401890.1401969>.
49. Anderson, A. C. et al. In-vivo shift of the microbiota in oral biofilm in response to frequent sucrose consumption. *Sci. Rep.* **8**, 14202 (2018).

50. Schoilew, K. et al. Bacterial biofilm composition in healthy subjects with and without caries experience. *J. Oral. Microbiol.* **11**, 1633194 (2019).
51. Thomas, A. M. et al. Alcohol and tobacco consumption affects bacterial richness in oral cavity mucosa biofilms. *BMC Microbiol.* **14**, 250 (2014).
52. Baker, J. L., Bor, B., Agnello, M., Shi, W. & He, X. Ecology of the Oral Microbiome: Beyond Bacteria. *Trends Microbiol.* **25**, 362–374 (2017).
53. Nobbs, A. H. & Jenkinson, H. F. Interkingdom networking within the oral microbiome. *Microbes Infect.* **17**, 484–492 (2015).
54. Fontaine, L. et al. A Novel Pheromone Quorum-Sensing System Controls the Development of Natural Competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.* **192**, 1444–1454 (2010).
55. Gardan, R., Besset, C., Guillot, A., Gitton, C. & Monnet, V. The Oligopeptide Transport System Is Essential for the Development of Natural Competence in *Streptococcus thermophilus* Strain LMD-9. *J. Bacteriol.* **191**, 4647–4655 (2009).
56. Hammer, B. K. & Bassler, B. L. Quorum sensing controls biofilm formation in *Vibrio cholerae*. *Mol. Microbiol.* **50**, 101–104 (2003).
57. Kong, K.-F., Vuong, C. & Otto, M. Staphylococcus quorum sensing in biofilm formation and infection. *Int. J. Med. Microbiol.* **296**, 133–139 (2006).
58. Preda, V. G. & Săndulescu, O. Communication is the key: biofilms, quorum sensing, formation and prevention. *Discoveries* **7**, (2019).
59. Chen, M. et al. Oxidative stress-related biomarkers in saliva and gingival crevicular fluid associated with chronic periodontitis: A systematic review and meta-analysis. *J. Clin. Periodontol.* **46**, 608–622 (2019).
60. Loesche, W. J. Oxygen Sensitivity of Various Anaerobic Bacteria. *Appl. Microbiol.* **18**, 723–727 (1969).
61. D'souza, L. L., Lawande, S. A., Samuel, J. & Pinto, M. J. W. Effect of salivary urea, pH and ureolytic microflora on dental calculus formation and its correlation with periodontal status. *J. Oral. Biol. Craniofacial Res.* **13**, 8–12 (2023).
62. Gaál Kovalčíková, A. et al. Urea and creatinine levels in saliva of patients with and without periodontitis. *Eur. J. Oral. Sci.* **127**, 417–424 (2019).
63. Nascimento, M. M., Gordan, V. V., Garvan, C. W., Browngardt, C. M. & Burne, R. A. Correlations of oral bacterial arginine and urea catabolism with caries experience. *Oral. Microbiol. Immunol.* **24**, 89–95 (2009).
64. Osmani, F. Can the salivary urea and stimulated saliva concentration be a marker of periodontal diseases in opioid users? A case-control study. *Heliyon* **9**, (2023).
65. Bentley, C. D. & Disney, J. A. A comparison of partial and full mouth scoring of plaque and gingivitis in oral hygiene studies. *J. Clin. Periodontol.* **22**, 131–135 (1995).
66. Heinrich-Weltzien, R., Kühnisch, J., Van Der Veen, M., De Josselin De Jong, E. & Stöber, L. Quantitative light-induced fluorescence (QLF) – A potential method for the dental practitioner. *Quintessence Int* **34**, 181–188 (2003).
67. Volgenant, C. M. C. et al. Red fluorescence of dental plaque in children – A cross-sectional study. *J. Dent.* **58**, 40–47 (2017).
68. Zaura, E. et al. On the ecosystemic network of saliva in healthy young adults. *ISME J.* **11**, 1218–1231 (2017).
69. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
70. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
71. Evans, A. M., DeHaven, C. D., Barrett, T., Mitchell, M. & Milgram, E. Integrated, Nontargeted Ultrahigh Performance Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry Platform for the Identification and Relative Quantification of the Small-Molecule Complement of Biological Systems. *Anal. Chem.* **81**, 6656–6667 (2009).
72. Lawton, K. A. et al. Analysis of the adult human plasma metabolome. *Pharmacogenomics* **9**, 383–397 (2008).
73. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551 (2021).
74. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
75. Silness, J. & Løe, H. Periodontal Disease in Pregnancy II. Correlation Between Oral Hygiene and Periodontal Condition. *Acta Odontol. Scand.* **22**, 121–135 (1964).
76. Van der Weijden, G. A., Timmerman, M. F., Nijboer, A., Lie, M. A. & Van der Velden, U. A comparative study of electric toothbrushes for the effectiveness of plaque removal in relation to toothbrushing duration: Timerstudy. *J. Clin. Periodontol.* **20**, 476–481 (1993).
77. Aitchison, J. The Statistical Analysis of Compositional Data. *J. R. Stat. Soc. Ser. B Methodol.* **44**, 139–160 (1982).
78. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* **8**, 2224 (2017).
79. Bro, R. & Smilde, A. K. Centering and scaling in component analysis. *J. Chemom.* **17**, 16–33 (2003).
80. Chen, T. et al. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* **2010**, (2010).
81. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
82. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
83. Kawashima, S., Katayama, T., Sato, Y. & Kanehisa, M. KEGG API: A Web Service Using SOAP/WSO to Access the KEGG System. *Genome Inf.* **14**, 673–674 (2003).
84. Martino, C. et al. Context-aware dimensionality reduction deconvolutes gut microbial community dynamics. *Nat. Biotechnol.* **39**, 165–168 (2021).
85. van der Ploeg, G. R., Westerhuis, J. A., Heintz-Buschart, A. & Smilde, A. K. parafac4microbiome: Exploratory analysis of longitudinal microbiome data using Parallel Factor Analysis. Preprint at <https://doi.org/10.1101/2024.05.02.592191> (2024).
86. Lorho, G., Westad, F. & Bro, R. Generalized correlation loadings: Extending correlation loadings to congruence and to multi-way models. *Chemom. Intell. Lab. Syst.* **84**, 119–125 (2006).
87. Borchers, H. W. & Borchers, M. H. W. Package 'pracma'. Accessed On 4, (2022).
88. Schmidt, E. Über die Auflösung Linearer Gleichungen mit Unendlich Vielen Unbekannten. in *Integralgleichungen und Gleichungen mit unendlich vielen Unbekannten* (ed. Pietsch, A.) 11 249–278 (Springer Vienna, Vienna, 1989).
89. Kiers, H. A. L. Some procedures for displaying results from three-way methods. *J. Chemom.* **14**, 151–170 (2000).
90. Maechler, M. Finding groups in data: Cluster analysis extended Rousseeuw et al. *R. Package Version* **2**, 242–248 (2019).
91. Kassambara, A. & Mundt, F. Package 'factoextra'. *Extr. Vis. Results Multivar. Data Anal.* **76**, (2017).

Acknowledgements

We want to thank Michelle van der Wurff (TNO) and Tim van den Broek (TNO) for their help in data acquisition, processing and storing and Jesse Alderliesten (UU/UvA), Fred White (UvA), Cynthia Albracht (UvA), Juan Pablo Bascur (CWTS) and Rianne Warmerdam (Aiden) for their useful discussions and suggestions. The authors thank N.A.M. Rosema for coordinating the clinical study. For assistance in capturing fluorescence photographs, the authors thank Y. Altındağ. For the clinical examinations the authors thank J.M. Voll (clinical plaque assessment with Silness & Løe), and S. Bizzarro (assessment of the bleeding on marginal probing). This research is

supported by the Dutch Technology Foundation STW (project number 10948) and the Top Institute Food and Nutrition (TIFN), a public-private partnership on precompetitive research in food and nutrition. Organizations supporting this project had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. QLF data was obtained with funding from NWO (ZonMw-STW-NIG-program): “Project 10948: Seeing is believing? A novel tool for the visualization of oral disease manifestations.” GRvdP was funded by a grant from the University of Amsterdam, Research Priority Area on Personal Microbiome Health.

Author contributions

G.R. van der Ploeg: formal analysis, methodology, writing—original draft; B.W. Brandt: investigation, writing—review and editing; B.J.F. Keijser: investigation, project administration, writing—review and editing; M.H. van der Veen: investigation, project administration, writing—review and editing; C.M.C. Volgenant: investigation, writing—review and editing; E. Zaura: conceptualization, investigation, project administration, funding acquisition, writing—review and editing; A.K. Smilde: conceptualization, methodology, supervision, funding acquisition, writing—review and editing; J.A. Westerhuis: conceptualization, methodology, supervision, formal analysis, writing—original draft; A. Heintz-Buschart: conceptualization, formal analysis, supervision, project administration, writing—original draft. All authors reviewed and approved the manuscript.

Competing interests

The authors declare no competing interests.

Ethics approval

The study involving human participants was conducted in accordance with the ethical principles of the 64th WMA Declaration of Helsinki (October 2013, Brazil) and the Medical Research Involving Human Subjects Act (WMO), approximating Good clinical Practice (CPMP/ICH/135/95) guidelines. The clinical trial was approved by the Medical Ethical Committee of the VU Medical Center (2014.505) and registered at the public trial register of the

Central Committee on Research Involving Human Subjects (CCMO) under number NL51111.029.14.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41522-024-00565-x>.

Correspondence and requests for materials should be addressed to A. Heintz-Buschart.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024