

## Article

# A Log Ratio-Based Analysis of Individual Changes in the Composition of the Oral Microbiota in Different Dietary Phases

Kirstin Vach <sup>1,2,\*</sup> , Ali Al-Ahmad <sup>2</sup> , Annette Anderson <sup>2</sup> , Johan Peter Woelber <sup>2</sup> , Lamprini Karygianni <sup>3</sup> , Annette Wittmer <sup>4</sup> and Elmar Hellwig <sup>2</sup>

<sup>1</sup> Institute of Medical Biometry and Medical Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Stefan-Meier-Str. 26, D-79104 Freiburg, Germany

<sup>2</sup> Center for Dental Medicine, Department of Operative Dentistry and Periodontology, Faculty of Medicine and Medical Center, University of Freiburg, Hugstetter Straße 55, D-79106 Freiburg, Germany; ali.al-ahmad@uniklinik-freiburg.de (A.A.-A.); annette.anderson@uniklinik-freiburg.de (A.A.); johan.woelber@uniklinik-freiburg.de (J.P.W.); elmar.hellwig@uniklinik-freiburg.de (E.H.)

<sup>3</sup> Clinic for Conservative and Preventive Dentistry, Center of Dental Medicine, University of Zurich, Plattenstrasse 11, CH-8032 Zurich, Switzerland; Lamprini.Karygianni@zzm.uzh.ch

<sup>4</sup> Institute of Medical Microbiology and Hygiene, Department of Medical Microbiology and Hygiene, Faculty of Medicine and Medical Center, University of Freiburg, Hermann-Herder-Straße 11, D-79104 Freiburg, Germany; annette.wittmer@uniklinik-freiburg.de

\* Correspondence: kv@imbi.uni-freiburg.de; Tel.: +49-761-203-5004



**Citation:** Vach, K.; Al-Ahmad, A.; Anderson, A.; Woelber, J.P.; Karygianni, L.; Wittmer, A.; Hellwig, E. A Log Ratio-Based Analysis of Individual Changes in the Composition of the Oral Microbiota in Different Dietary Phases. *Nutrients* **2021**, *13*, 793. <https://doi.org/10.3390/nu13030793>

Academic Editor: Eva Untersmayr

Received: 1 February 2021

Accepted: 23 February 2021

Published: 28 February 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Background: Investigating the influence of nutrition on oral health has a long scientific history. Due to recent technical advances like sequencing techniques for the oral microbiota, this topic has gained scientific interest again. A basic challenge is to understand the influence of nutrition on the oral microbiota and on the interaction between the oral bacteria, which is also statistically challenging. Methods: Log-transformed ratios of two bacteria concentrations are introduced as the basic analytic tool. The framework is illustrated by application in an experimental study exposing eleven participants to different nutrition schemes in five consecutive phases. Results: The method could be sufficiently used to analyse the interrelation between the bacteria and to identify some bacterial groups with the same as well as different reactions to additional dietary components. It was found that the strongest changes in bacterial concentrations were achieved by the additional consumption of dairy products. Conclusion: A log ratio-based analysis offers insights into the relation of different bacteria while taking specific features of compositional data into account. The presented methods allow becoming independent of the behaviour of other bacteria, which is a disadvantage of common analysis methods of compositions. The results indicate that modulations of the oral biofilm microbiota due to nutrition change can be attained.

**Keywords:** compositional data; ratio fractions; nutrition; microbiome; oral health

## 1. Introduction

In recent years, the influence of nutrition on oral biofilm has received increasing attention. Dietary factors have a major influence on the microbiota of the oral biofilm and can lead to disturbances of the homeostasis of the more than 700 bacterial species within its complex network. For example, frequent consumption of simple carbohydrates can favour acidogenic and aciduric bacterial species, consequently leading to carious lesions [1]. On the other hand, frequent consumption of dairy products, dietary fibre and certain vegetable foods is associated with lower caries incidence [2–4]. The calcium, phosphate and casein phosphopeptides contained in dairy products can counteract acidic demineralisation and promote remineralisation processes, thus having a caries preventive effect. In terms of plant-based foods, different ingredients are concerned with either promoting remineralisation (e.g., phosphates) or reducing bacterial adhesion and growth, thereby reducing the risk of caries development [1,5–7]. It is known that simple carbohydrates can lead to a higher

proportion of acidogenic and aciduric species, which can promote the development of carious lesions [8]. Concerning the different bacterial species involved in the caries process, for a long time, mutans streptococci and lactobacilli were thought to be the main species responsible for the acid production and demineralisation of the tooth structure. However, in recent years it has become clear that various additional bacterial species are associated with different stages of caries development: non-mutans streptococci, such as *Streptococcus salivarius* and *Streptococcus parasanguinis*, and species of the genus *Actinomyces* have been associated with the initial stages, while *Veillonella* spp., *Lactobacillus* spp., *Atopobium* spp. and other taxa dominate the more advanced stages [9–11]. The effect of nutrition on the oral microbiota has mostly been investigated with either in vitro experimental approaches, testing individual ingredients, using animal experiments or analysing epidemiological data. In this paper, we consider a study analysing the influence of certain foods on the whole microbiota of the oral biofilm, was tested in situ using splint systems worn by study participants on which the oral biofilm was grown. The bacterial composition of the oral biofilm was then analysed by culture technique [1].

When analysing this type of oral biofilm data, it is common practice to examine the fractions of total concentration rather than the pure concentrations. These fractions sum up to 1, which causes special challenges in terms of the statistical analysis. Data of this kind are called compositional data, and its analysis has attracted strong attention in recent years [12–16]. We present a case study, in which we apply a log ratio-based approach to oral biofilm data. This approach allows analysing changes in the composition when additional standardised dietary intakes are available and the interest lies in comparing with one baseline condition (regular diet).

## 2. Materials and Methods

### 2.1. The Data

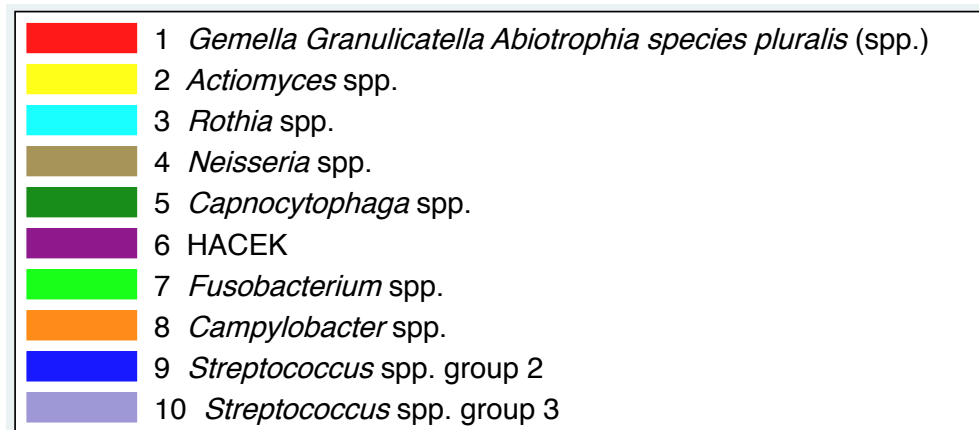
We consider a study funded by the German Research Foundation (DFG) with eleven participants and data investigating the influence of additional standardised diet components on the microbiota of the oral biofilm over 15 months. The participants underwent five phases, each of which lasted three months. During these phases, standardised dietary components in the form of sucrose, dairy products and vegetables in addition to the regular diet were investigated. Phase I was a lead-in phase without changing the nutrition habits. In phase II, participants consumed an additional 10 g of rock candy per day (Weisser Kandis; Südzucker AG, Mannheim, Germany), sucking small pieces of 2 g five times in between meals. In phase III, additional milk products were consumed (150 g of plain yoghurt 3 times daily and 100 mL of long-life milk twice a day, both 1.5% fat; Schwarzwaldmilch GmbH, Freiburg, Germany). In phase IV, 500 g of vegetable puree (types: white carrot, parsnip, carrot, “jardinière” (carrot, potato, cauliflower and pea), pumpkin and garden vegetables (carrot, potato, spinach, parsnip and leeks), Reine Weiße Karotte, Reine Pastinake, Reine Frühkarotte, and Gemüse-Allerlei (Hipp GmbH, Pfaffenhofen, Germany) and Kürbis pur and Gartengemüse (Alnatura, Darmstadt, Germany)) per day was given, before the participants returned to their individual regular diet in phase V.

An in situ splint system was used to sample the dental plaque. Individual upper jaw rigid acrylic appliances were manufactured for each study participant [17]. The splint system was taken out and stored in saline solution (0.9% NaCl) for regular meals and during oral hygiene only. It was worn while consuming the phase-specific additional foods, which were eaten slowly to expose them to the oral cavity for several minutes. During all phases, the oral biofilm was allowed to grow on embedded enamel slabs over the course of seven days. Subsequently, the splint system was removed for analysis of the oral biofilm, cleaned and after seven days it was reapplied for another seven days. This procedure was repeated three times all together. Participants brushed their teeth with standardised tooth brushes and toothpaste with a sodium fluoride content of 1450 ppm.

The oral biofilm grown in all five phases was analysed using culture techniques averaging over three measurements, which were taken within four weeks per phase.

Twenty-seven groups of bacteria were identified from a total of 93 different cultivated species. Details about the course of the study and data collection can be found in [18,19].

A main focus of this work was to ascertain whether the additional standardised dietary intake leads to changes in oral microbiota. It makes sense to only consider bacteria that are usually above the detection limit. According to the results of [19], only disjunct bacterial groups were chosen, where the percentage of values above the detection limit was greater than 75%. All values below the detection limit were set to the detection limit. The bacterial groups together with the colour scheme that we use in all of our figures can be found in Figure 1. A detailed list can be found in the Appendices A and B.



**Figure 1.** Bacterial groups and colour scheme for all figures.

If the sum of the concentrations of the bacterial groups listed here is used as the total concentration, the fractions of the single bacterial groups sum up to 1 and the data has a compositional character. For each bacterial group, participant and phase, the starting point is considered to be its fraction among the total concentration of the analysed bacterial groups.

## 2.2. Analytic Strategy

- For each pair of bacteria, the ratio between the two fractions is built for each participant and phase. Such a pairwise ratio directly depicts the relation between two bacteria. They do not depend on the concentration of the other bacteria or the overall concentration. If a third bacterium has an unusually high or low concentration, this affects the overall concentration and hence the fractions, but it does not affect the ratio. Therefore, in this way we become independent of the behaviour of other bacteria. Ratios also have a simple direct interpretation describing how more frequent one bacterium is compared with the other. Note that a ratio between two fractions can also be interpreted directly as the ratio between two corresponding concentrations.
- The second step is to consider the logarithm with the base 10 of the ratio. This is useful because ratios tend to have a skewed and unstable distribution, whereas taking logarithms leads to a more symmetric distribution of the values. The basis 10 was chosen due to its strong interpretability: a value of 1 in the  $\log_{10}$  scale means the tenfold in the real concentration setting. A ratio of 1—i.e., equal concentrations—has a value of 0 in the  $\log_{10}$  setting. For ratios smaller than 1, the  $\log_{10}$  transformed value becomes negative. Therefore, a positive (negative) value means that the first bacteria is less (more) frequent than the second one.
- Logarithmising the ratios leads to one number per participant and phase, which can be analysed just like any other number [15]. In particular, mean values and standard deviations can be calculated. In this step, for each pair of bacteria we consider the mean value of the log ratios across all participants per phase. It is interpretable as the typical relationship between these two bacteria in this phase. Therefore, this

mean value offers the opportunity to check, whether the relationship between two bacteria is similar in different phases or whether it changes due to changes in the nutrition. Accordingly, we can analyse the stability between bacteria between the phases. Similarly as above, a mean value of 1 suggests that typically the first bacterium has ten times the concentration of the second bacterium.

4. We also analyse the corresponding standard deviations, which are a measure of the stability of the relationship between two bacteria within one phase. The closer that the value is to 0, the more homogeneous are the values across participants, i.e., the relation between the two considered bacteria is stable and can thus be assumed as “typical”. For example, if we observe a standard deviation of 0.5 and a mean log ratio of 0 for two bacteria, then the application of the 2 sigma rule leads to a 95% range of (−1, 1), i.e., 95% of the observed ratios are between 1/10 and 10.
5. As the aim is to study the impact of an additional standardised diet on the microbiome, we also depict the changes of the concentrations compared with the first phase. For a given pair of bacteria, for each participant and phase we consider the difference in the log ratio compared with the first phase. A difference above 0 means that the ratio increased, i.e., the first bacterium became more frequent relative to the second one. As a difference in the log ratio is equal to the logarithm of the ratio of ratios, we can also interpret the values themselves. A difference of 1 means, that the ratio has increased by a factor of ten. For example, this may reflect a change from 1.2 to 12, or from 0.3 to 3. Again, we can consider mean values and standard deviations across participants per phase for phases II to V. The mean values inform us about the typical change in the relation between the two bacteria. The standard deviations offer a hint about the stability of these changes across the participants.
6. Finally, we also consider a bacterium-specific summary measure, which is based on averaging the log ratio changes for one bacterium over all other bacteria within a single participant. If this way a bacterium has a value greater than 0, this can be interpreted as having increased its fraction on average relative to all other bacteria. The bacterium with the largest positive value is the one with the most distinct increase compared with all other bacteria. We refer to this value for each participant as the average log ratio change.

A detailed mathematical description of the different quantities can be found in the Appendix B. For all means and standard deviations, 95% confidence intervals can be computed using statistical standard methods. We only report them for the most fundamental analyses, focussing on the mean values in pairs and the mean average log ratios. For the analyses, the statistics program STATA (StataCorp LT, College Station, TX, USA, Version 16.1) was used. For graphical presentation, bar and pie charts, scatter plots and forest plots were used.

### 2.3. Analysing Change or Comparing Phases Directly

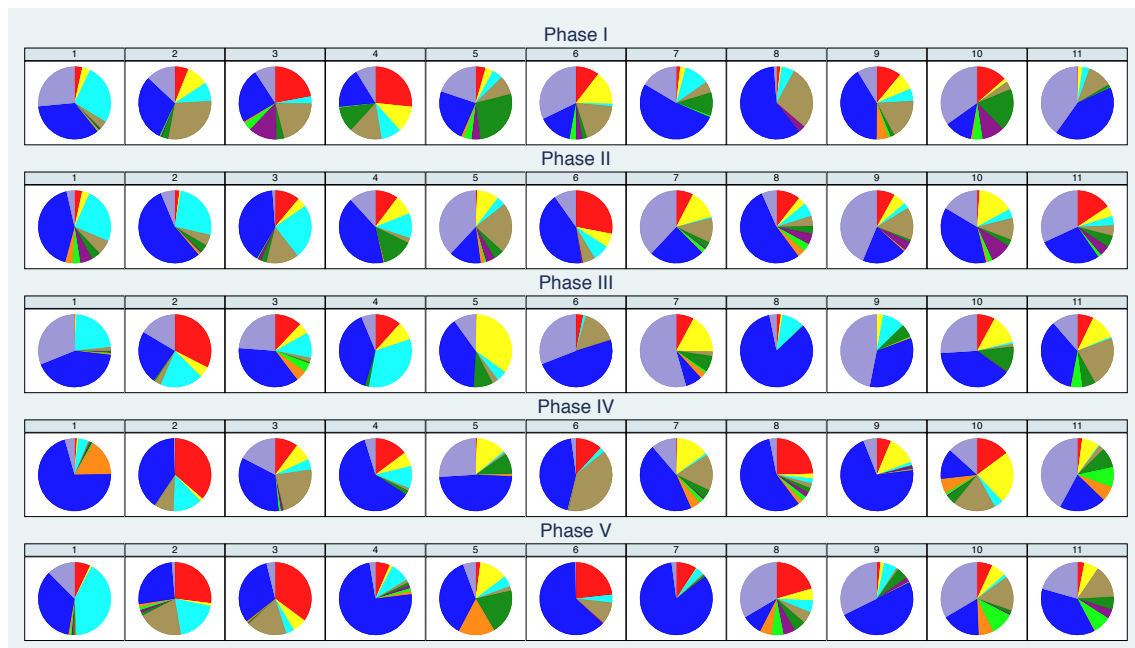
In studies using a baseline condition, it is usual to consider the individual changes from the baseline, in order to investigate the impact of the subsequent conditions. However, this is not necessarily the most adequate strategy in the context of this study. It may be the case that the standardised dietary additions imply that all participants exhibit a common composition in each phase, overruling the individual differences at the baseline. In such a scenario, it would be more adequate to compare the composition directly between the conditions instead of investigating the change from the baseline. The standard deviations of the log ratios will inform us about the homogeneity of bacteria spectra in each phase, whereby we will be able to judge whether the scenario is present or not.

## 3. Results

### 3.1. A First Look at the Raw Data

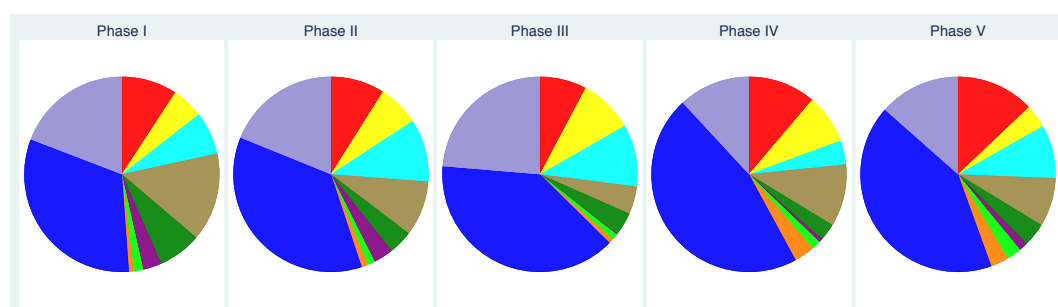
In Figure 2, the fractions of the total concentrations for all ten bacterial groups are shown for each participant for each of the five phases. If the nutrition leads to a certain

bacteria pattern, the pictures would have to be similar in one line each, while they would differ in the columns. This should at least apply from phase II to phase IV, as a change to a specific additional diet took place from this point onwards. However, instead of uniform patterns in each phase, we observe rather individual pattern across phases, similar to phase I. Here, the different patterns are unsurprising for the different participants, because the nutrition was in no way influenced.



**Figure 2.** Fractions of the bacterial groups on the total concentration per participant (column) and phase (row) based on mean values per phase.

If we now consider the mean values per phase (Figure 3), we see that specific bacterial groups are actually rather common in all phases. We also observe that the differences between the phases are quite small.

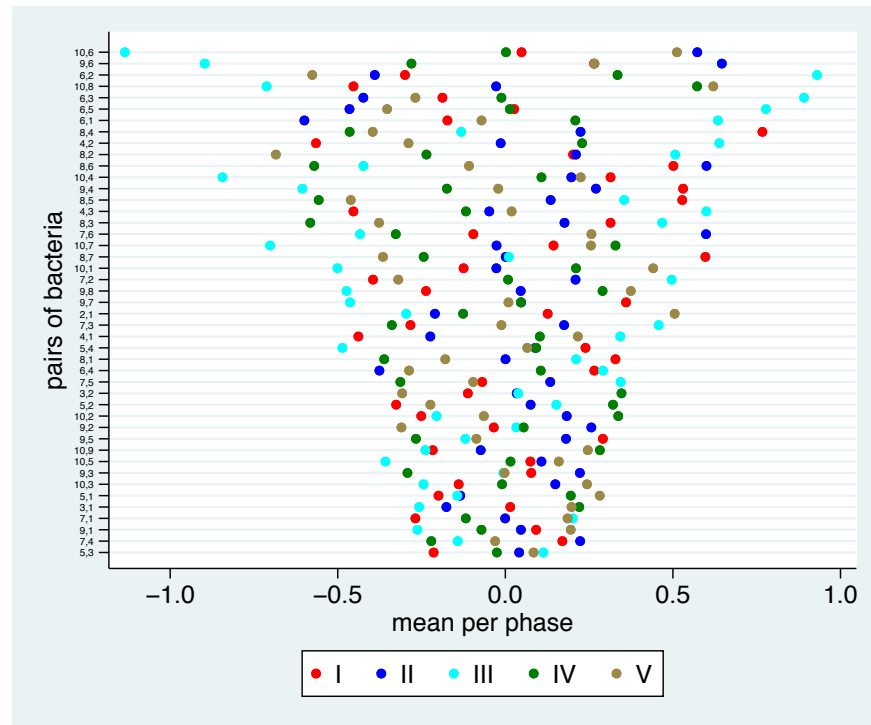


**Figure 3.** Fractions of the bacterial groups on the total concentration per phase based on mean values per phase.

### 3.2. Stability of the Relation between Bacterial Groups between Phases

The relation between two bacterial groups may remain stable between the phases, or it may change. Stability may indicate that the relation is minimally affected by changes in nutrition, whereas instability may indicate sensitivity to changes. For each combination of two bacterial groups, we look at the log ratio for each participant in each phase and consider the phase-specific mean value. We can see in Figure 4 that there are differences between the pairs. At the bottom of the picture, we find pairs, where the relation to each other is very similar over all phases. At the top of the picture, we see pairs with distinct differences across the phases. Note that among the pairs with the lowest variation in the mean log ratio, we find bacterial group numbers 1, 3, 5 and 7 several times. In four instances

we find bacterial group 1, and twice we find bacterial groups 3, 5 and 7. In particular, all combinations of the bacterial groups 1, 3 and 5 are among the pairs with the lowest variation in the mean log ratio.



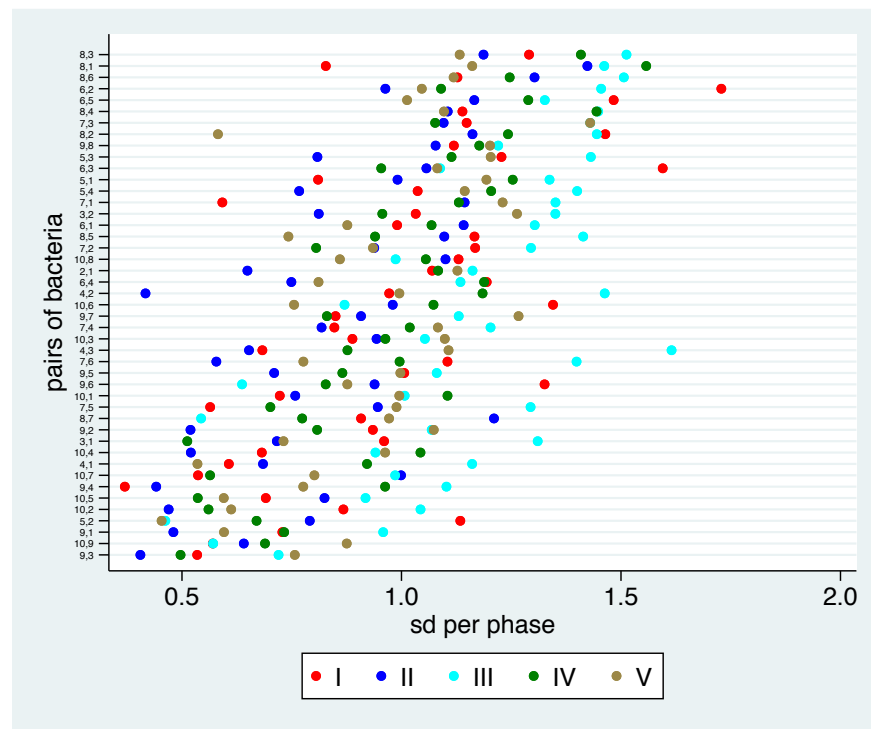
**Figure 4.** Mean log ratios for all bacterial pairs of bacterial groups and phases. The values are centred around 0 for each bacterial pair. The bacterial pairs are ordered according to the range of the mean values.

Further, the mean log ratio in phase III is often far on the left or far on the right, i.e., in this phase we see the most extreme ratios.

### 3.3. Stability of the Relation between Bacterial Groups within Phases

The mean values above offer a hint whether the pairwise relation between bacterial groups is common in different phases. Now the question remains how stable the pairwise relation between bacterial groups is within one phase. Therefore, we take a look at the standard deviations of the phase-specific log ratios in Figure 5.

Some bacterial pairs show low values for all phases, such as 9-3, 9-10, 9-1, 5-2, 10-2 and 10-5. This means that the relation between these bacteria is stable within participants within each phase. However, even for these pairs we observe standard deviations in the magnitude of 0.5 or above, indicating rather wide 95% ranges across participants. This corresponds with the high individual variation observed in Figure 2. It is interesting to note that all six of these bacterial pairs also showed a low variation in mean values across the phases. They were all among the fourteen bacterial pairs (out of 45) with lowest range in Figure 4, i.e., these bacterial pairs tend to be stable in their relation across participants *and* phases. The opposite relation does not seem to hold: the bacterial pairs 5-3, 7-1 and 5-1 show little variation in mean values, but nevertheless high standard deviations within each phase. For some bacterial pairs, the stability substantially varies across the phases. For example, the bacterial pairs 8-1 and 7-1 show high stability in phase I and low stability in other phases, suggesting that the specific dietary supplements destroy a rather stable relation seen at the baseline. Another example is the bacterial pair 4-3, where phase III shows a much higher standard deviation than the other phases.



**Figure 5.** Standard deviations (sd) for phase-specific log ratios for all bacterial pairs of bacterial groups and phases based on empirical estimates. The pairs are sorted by the mean sd value over all phases.

If we compare the phases I–IV (the nutrition in phase V was analogous to phase I by study design), the lowest standard deviation can be found in twelve out of 45 bacterial pairs in phase I, suggesting that there is no tendency towards a more uniform composition after standardised dietary supplementation. Therefore, there is no argument against analysing the changes, i.e., the changes from the baseline.

### 3.4. Identification of Bacterial Pairs with a Change in Relation in Comparison with the Baseline

In Figure 6, the mean log ratio changes with 95% confidence intervals for all pairs of bacteria are shown. While in phase II the additional diet only leads to few changes, more reactions can be observed in the other phases. The observed patterns are very similar in phases IV and V. Overall, we have to conclude that definitive statements about a “statistically significant” change in the relation in one specific phase are difficult to justify, as only few of the overall 180 confidence intervals do not include 0. However, in phase III eleven, in phase IV seven and in phase V six out of 45 confidence intervals do not include 0, i.e., we observe more than the 2.25 that we should expect under chance level. Consequently, the observed mean log ratio changes do not simply reflect random noise.

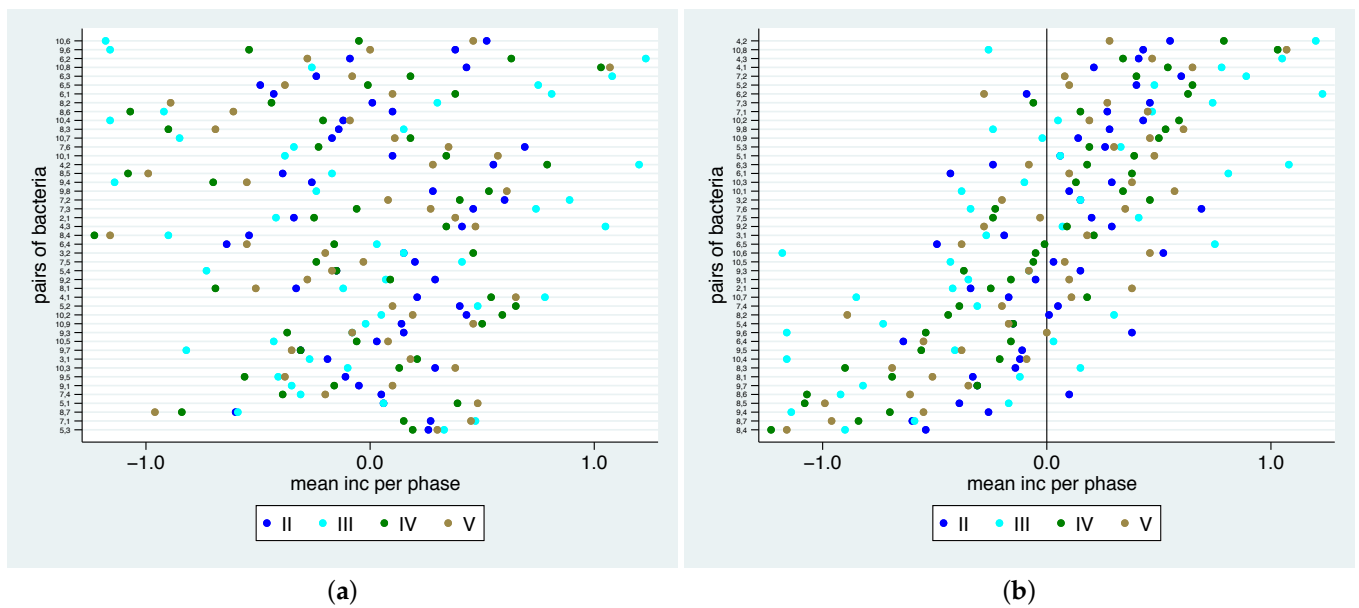


Figure 6. Mean log ratio changes for all bacterial pairs for phases II–V with 95% confidence intervals.

### 3.5. Stability of Change in Relation between Phases

To gain a better overview of these numbers and in particular the relation to the different phases, in Figure 7 we show the phase-specific mean values for each bacterial pair. We have sorted the pairs of bacteria by (a) their range within the different phases (small ranges on the bottom, larger ranges on the top) and (b) by the mean over the phases. In Figure 7a, we observe that the bacterial pairs 5-3, 7-1, 8-7, 5-1, 7-4, 9-1 and 9-5 show the smallest range, meaning that they show a common reaction on the additional diet in all phases, while the bacterial pairs 10-6, 9-6, 6-2, 10-8, 6-3, 6-5 and 6-1 show a large variation in their reactions in different phases. Additionally, one can see in Figure 7b that the bacterial pairs 4-2, 4-3, 4-1, 7-2, 5-2, 7-1, 10-2, 5-3 and 5-1 (8-4, 8-7, 9-4, 8-5, 9-7, 8-1, 10-4 and 9-5) show a positive (negative) value in all phases, meaning that the second bacterium tends to be more (less) frequent than the first one in each phase. Several bacterial pairs show relation changes into different directions in different phases, which could be a sign that they react differently to the nutrition change.

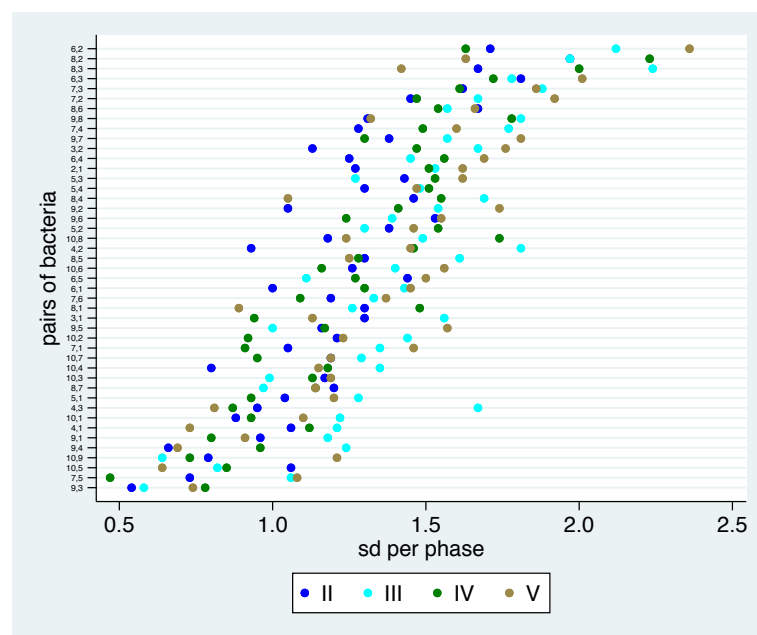




**Figure 7.** Mean values for phase-based log ratios for changes for all pairs of bacteria and phases sorted by range (a) and mean (b).

### 3.6. Stability of Change in Relation within Phases

The corresponding estimated standard deviations are shown in Figure 8. We observe some bacterial pairs, where the reaction to the additional diet is rather uniform across participants in all phases. These are nearly the same bacterial pairs such as in Figure 5, namely, 9-3, 7-5, 10-9, 10-5, 9-4, 9-1, 4-1 and 10-1. The bacterial pairs that always show high instability—6-2, 8-2, 8-3, 6-3 and 7-3—are also nearly identical to the former observations. Note that only the pairs 9-1, 9-5, 9-4 and 8-7 with low variation in mean values across the phases also show stability within phases. By contrast, we observe lower values in standard deviations for the bacterial pair 10-1, but high variation in mean values. The least uniform reaction often occurs in phase III.



**Figure 8.** Standard deviation (sd) for phase-based log ratios for changes for all pairs of bacteria and phases estimated empirical and sorted by mean of sd over the phases.

### 3.7. Interim Summary

In Table 1, we present an overview of the results of the four previous analyses: the bacterial pairs with the lowest or highest stability in the four different analyses are marked with letters differentiating between the analyses performed on the raw values (small letters) and on change values (capital letters) and between analyses within and between phases.

**Table 1.** Overview of results from the analyses in Sections 3.2–3.6: The ten bacterial pairs with the highest (half-matrix left bottom) and the lowest (half-matrix right top) stability in each analysis are marked with b (between) and w (within) phases and corresponding capital letters for changes in comparison to baseline.

	1	2	3	4	5	6	7	8	9	10
1						<i>bB</i>		<i>w</i>		
2				<i>b</i>		<i>bwBW</i>	<i>W</i>	<i>bwBW</i>		
3	<i>bB</i>				<i>w</i>	<i>bBW</i>	<i>wW</i>	<i>wW</i>		
4	<i>wW</i>		<i>W</i>				<i>W</i>	<i>bw</i>		<i>B</i>
5	<i>bBW</i>	<i>w</i>	<i>bB</i>	<i>W</i>		<i>bwB</i>				
6								<i>wBW</i>	<i>bB</i>	<i>bB</i>
7	<i>bB</i>			<i>bB</i>	<i>W</i>				<i>W</i>	
8							<i>B</i>		<i>wW</i>	<i>bB</i>
9	<i>bwBW</i>		<i>bwW</i>	<i>w</i>	<i>B</i>		<i>B</i>			
10	<i>W</i>	<i>w</i>	<i>bB</i>	<i>w</i>	<i>bwW</i>		<i>w</i>		<i>bwW</i>	

For example, the bacterial pair 3-1 is among the ten bacterial pairs with the highest stability with respect to the analysis of the pairwise relation between phases and the change in relation between phases, and thus it is marked as *bB*. The bacterial pair 4-8 is among the ten bacterial pairs with the lowest stability when analysing the pairwise bacterial relation between and within phases, and thus it is marked with *bw*. First of all, we can state that many pairs can be found more than once among those with the highest (or lowest) stability, indicating a rather robust classification as “stable” (or “unstable”). Sub-patterns that we typically observe are *bB* and *wW*, indicating that analysing raw values or changes provide similar results. This mainly reflects the fact that high stability in raw values also implies some stability in changes, and that low stability in raw values also makes a low stability in changes likely if the correlation between phases is only moderate.

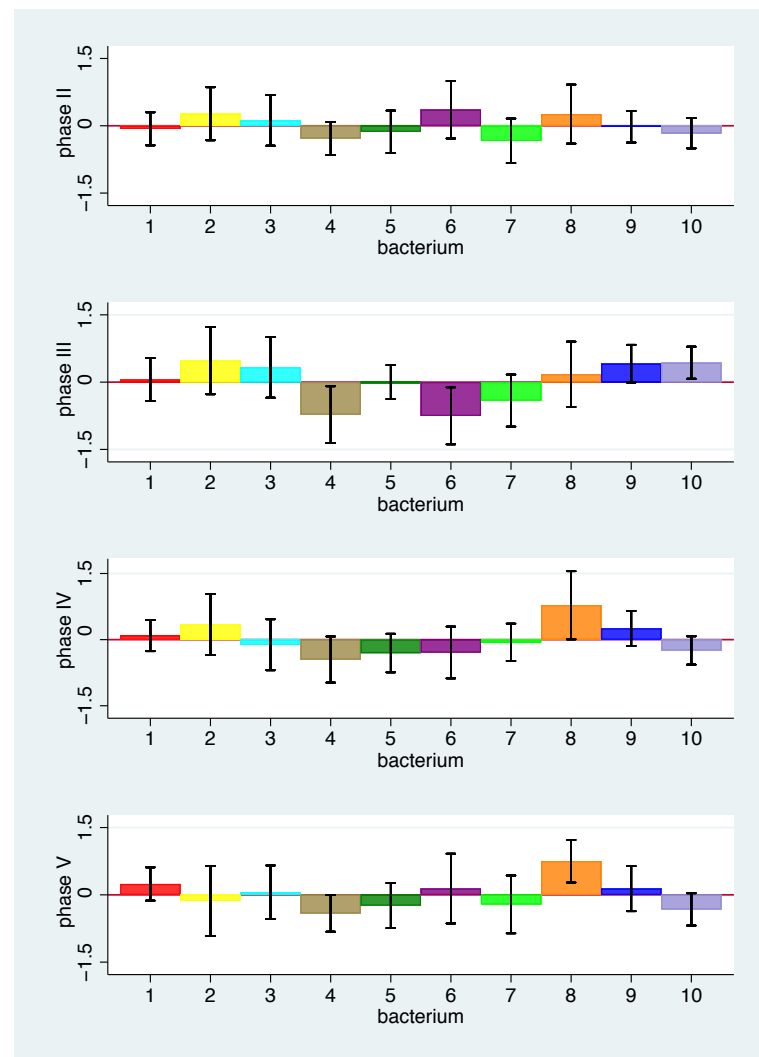
Conceptually, it is more interesting whether stability within phases coincides with stability between phases, i.e., the *bwBW* pattern in the lower triangle. We can identify one bacterial pair (1-9) with high stability across all analyses. The bacterial pairs 1-5, 3-9, 5-10 and 9-10 show high stability in at least both a within-phase and between-phase analysis of the same values (*bw* or *BW* sub-pattern in the lower triangle). The relation between these pair of bacteria seems to be difficult to modify by dietary supplements or other external factors. On the other hand, we can identify two bacterial pairs with low stability in all analyses: 2-6 and 2-8 (i.e., the *bwBW* pattern in the upper triangle) and the bacterial pairs 3-6, 4-8 and 6-8 with low stability in at least both a within-phase and between-phase analysis (*bw* or *BW* sub-pattern in the upper triangle). The relation between these pairs of bacteria seems to be highly modifiable by diet as well as other external factors.

Pairs with high stability within phases but low stability between phases might hold the strongest interest, as these are the pairs that most distinctly react to the dietary supplements. Indeed, we can find one of such pairs 4-10 (*w*, *W* or *wW* in the lower triangle and *b*, *B* or *bB* in the upper triangle). Interestingly, there are also three pairs with the opposite pattern: 3-5, 4-7 and 7-9 (*w*, *W* or *wW* in the upper triangle and *b*, *B* or *bB* in the lower triangle), i.e., high stability between phases but low stability within phases. Finally, note that bacterial groups 6 and 8 have an entry together with nearly all other bacterial groups in the upper triangle

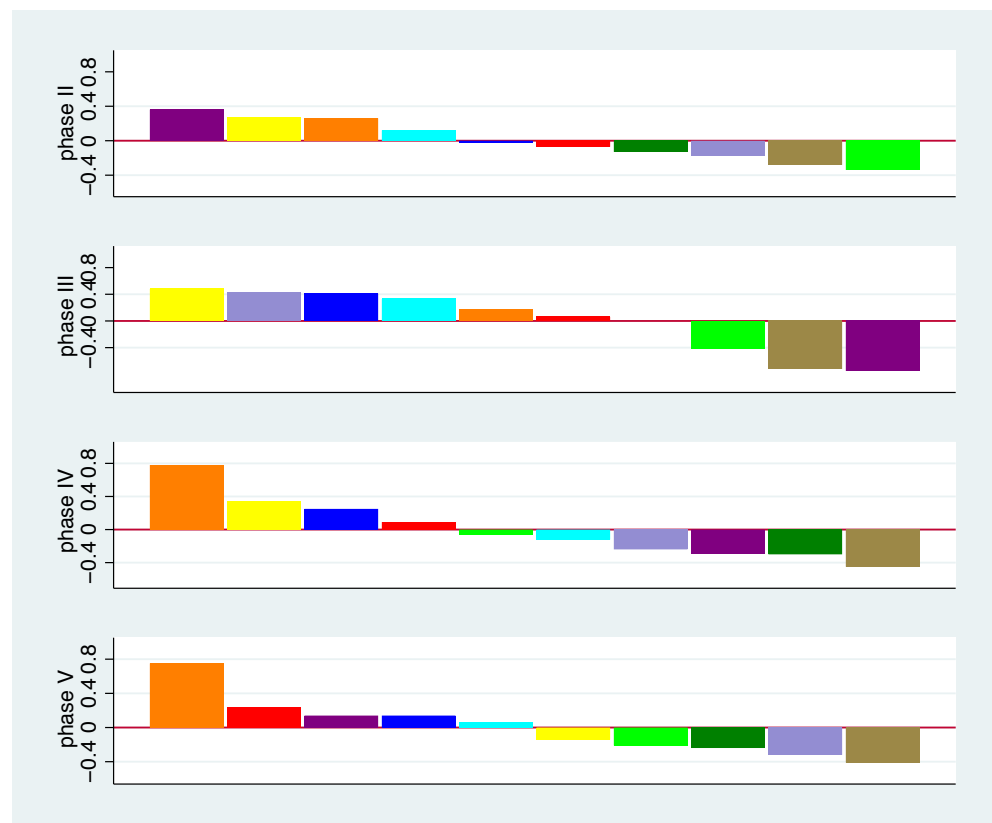
with low stability to other bacterial groups, while the bacterial group 1 is marked with most of the bacterial groups in the lower triangle with high stability. This may reflect the notion that bacterial group 1 is rather stable and bacterial groups 6 and 8 are rather unstable. Bacterial group 10 is striking: while it still showed little variation among the participants, one has a large variation between the phases.

### 3.8. Average Log Ratio Changes

Figure 9 shows the mean values of the average log ratio changes for each bacterial group in each phase. Definite statements about a specific change in a specific phase are difficult to make, as only few confidence intervals do not include 0. However, we observe that some bacterial groups tend to increase (bacterial group 8) or decrease in frequency (4, 7) for all phases, whereas other bacterial groups tend to differ in the direction of change over phases (e.g., 6, 10). We see the smallest overall changes in phase II. In Figure 10, within each phase we sort the bacterial groups by mean values of the average log changes to focus on the phase-specific pattern in the change of the bacterial spectrum. A striking feature here is the similarity of the pattern between phases IV and V. This is quite astonishing, as in phase V the same nutrition is given as with the baseline. One possible explanation would be that the effect of the nutrition change will last longer.



**Figure 9.** Mean values of average log ratio changes with 95% CI for bacterial groups 1 to 10 for phase II (top) to phase V (bottom). For colour scheme, see Figure 1.



**Figure 10.** Mean values of average log ratio changes sorted for bacterial groups 1 to 10 for phase II (top) to phase V (bottom). For colour scheme, see Figure 1.

### 3.8.1. Global $p$ -Value Per Phase

When looking at the single bacterial species, the question remains whether a global effect per phase can be observed, i.e., which additional diet led to a concentration change in nearly all bacterial groups. The global  $p$ -value (Table 2, the test used is explained in Appendix B.2) only shows a value that comes near to significance ( $p = 0.055$ ) for phase III.

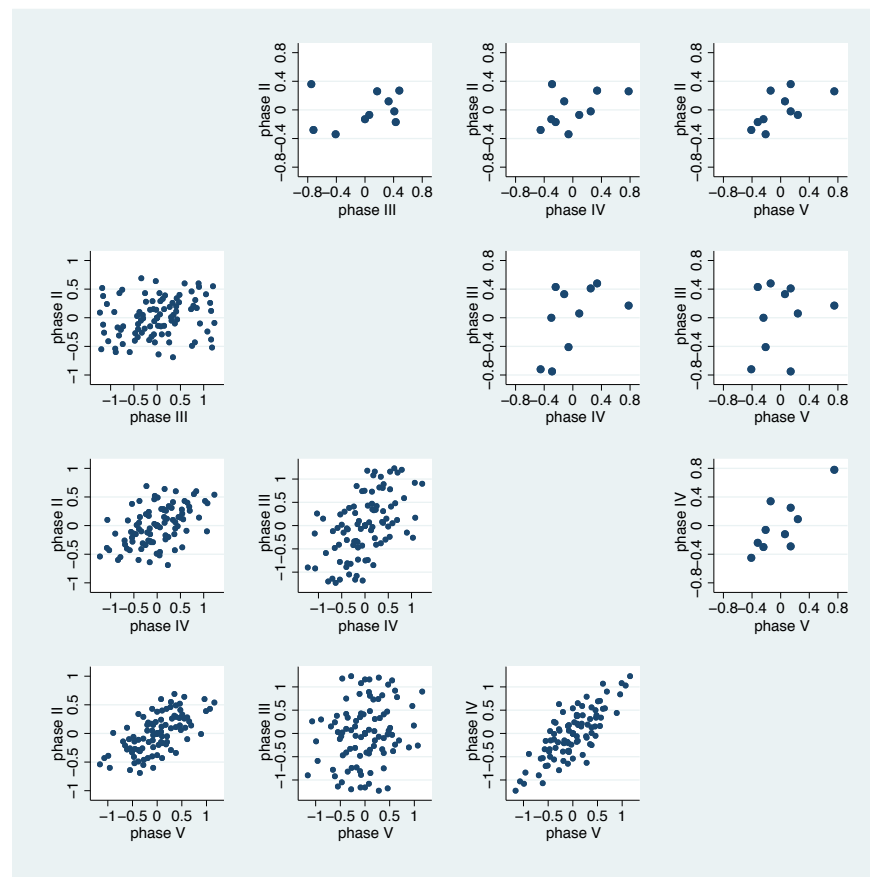
**Table 2.** Global  $p$ -value per phase.

	Phase II	Phase III	Phase IV	Phase V
$p$ -value	0.459	0.055	0.096	0.289

### 3.9. Similarity of Change Pattern Across Bacterial Groups

Figures 6 and 10 indicate some similarity in how bacterial groups react to the additional diet across different phases. We investigate this in more detail in Figure 11. We observe stronger correlations for phase II against phase V and phase IV against V. This impression is strengthened if we look in more detail at the left bottom, where all bacterial combinations are plotted.

The correlation coefficients (Table 3) confirm this impression.



**Figure 11.** Scatter plots of the mean values of bacterial groups' specific average log ratio changes. Half-matrix right top: mean value for each bacterial group in one phase against the other phases. Half-matrix left bottom: mean value for each bacterial pair in one phase against the other phases.

**Table 3.** Correlation coefficients for bacterial groups' specific average log ratio changes between phases.

	Phase III	Phase IV	Phase V
Phase II	0.199	0.488	0.628
Phase III		0.528	0.205
Phase IV			0.780

#### 4. Discussion

In this paper, we have presented an advantageous statistical approach to analyse a study investigating the influence of specific additional diets on the bacterial composition in the oral microbiota in an experimental manner. Traditionally, such analyses are focused on the (absolute or relative) concentration of the single bacterial species. Here, we adopt a different approach, looking at all pairs of bacteria and focusing on the quantitative relation of the two bacteria within each pair.

While this approach naturally adds complexity, it has one basic conceptual advantage that is particularly important when analysing a change in the spectrum: if the concentration or fraction of one bacterium increases or decreases distinctly, this does not directly affect the relation between any other pair of bacteria. Indeed, the simultaneous investigation of all bacterial pairs revealed that there are some pairs of bacteria that have a rather stable relation across participants within each phase. Some of them also have a rather stable relation across the phases, while some do not. Similar behaviour can be observed when examining the changes compared with the baseline.

It is rather natural that the simultaneous consideration of all bacterial pairs is not very useful to allow very specific statements for single pairs, in particular if the sample size is limited, as in our study. The general approach is mainly useful to draw an overall picture of bacterial correlation with the diet of the different phases of this longitudinal clinical study. For example, it became rather obvious that the diet of phase II had a lower impact on the composition than the dietary supplement in phases III and IV. Additionally, we could not observe the expected move back to the baseline scenario in phase V.

We could also address the fundamental question of whether an analysis of the study should focus on the individual changes in composition or on a typical composition for each nutrition. We found no evidence for the latter, and thus could stick to the intended analysis of the changes relative to the baseline.

With respect to the concrete insights gained by our analysis, the main results present the following points:

1. The bacterial pair *Gemella Granulicatella A. spp.*–*Streptococcus spp.* group 2 is the one with high stability across all analyses.
2. The bacterial pairs *Actiomyces. spp.*–HACEK and *Actiomyces. spp.*–*Campylobacter spp.* are the bacterial pairs with low stability across all analyses.
3. The bacterial pair *Neisseria spp.*–*Streptococcus spp.* group 3 could be identified as the bacterial pair that most distinctly reacts to the dietary supplements.
4. Bacterial pairs with high stability in at least both a within-phase and between-phase analysis were *Gemella Granulicatella A. spp.*–*Capnocytophaga spp.*, *Rothia spp.*–*Streptococcus spp.* group 2, *Capnocytophaga spp.*–*Streptococcus spp.* group 3 and *Streptococcus spp.* group 2–*Streptococcus spp.* group 3.
5. The most distinct changes in the spectrum were observed in phase III, i.e., the phase with additional milk products. In particular, the relation of *Neisseria spp.* to *Streptococcus spp.* group 2 and group 3 changed. When looking at the single bacterial species, the most distinct change was a decrease of *Neisseria spp.* and HACEK.
6. In phases IV and V, when looking at the single bacterial species, the increase of *Campylobacter spp.* should be mentioned. On the other hand, the figures in phases IV and V were very similar, although in phase V the participants returned to their individual nutrition without specific additions.
7. For *Gemella Granulicatella A. spp.*, *Rothia spp.* and *Streptococcus spp.* group 3, we observed rather stable relations to other bacterial groups, while for HACEK and *Campylobacter spp.* the relations are rather variable.

At least some of these findings can be explained. As *Gemella Granulicatella A. spp.* and *Streptococcus spp.* groups 2 and 3 are kinds of the *Streptococcus* species or relatives, they can all utilise the same substrate, and consequently a similar reaction is unsurprising. This explains our main result 1 and partly the results 4 and 7. HACEK are fastidious microorganisms, which could be an explanation for the fact that they also react quickly and strongly to changing environmental conditions, which supports the results 2, 5 and 7. The high variability of *Campylobacter spp.* (result 6 and 7) could be explained by the fact that in the presence of *Actiomyces spp.* and *Streptococcus spp.* the formate can be utilised by *Campylobacter spp.*, which *Actiomyces spp.* and *Streptococcus spp.* produce in glycolysis.

Note that though different microbial compositions were revealed in the different diet phases and mainly after the additional consumption of dairy products, the detected bacterial species have different potential for fermentation of carbohydrates by glycolysis leading to acids like formate, acetate and propionate. To avoid demineralisation of enamel and subsequently caries development a balanced diet is required which can sustain a balanced microbial composition of the supragingival oral biofilm.

#### Further Insights

In this type of experimental setup, it cannot be expected that the individual oral biofilm returns to its original composition in a short time when participants are allowed to return to their individual nutrition without studying specific additions. Inspection of the diaries

of the participants suggests that they did indeed return to the individual nutrition and did not continue to use higher levels of carbohydrates, fats and proteins. Further research including a combination of different microbiological analysing methods is necessary to understand whether this observation is indeed due to a long-lasting effect of nutrition on the oral biofilm. This would have a direct impact on the design of future studies.

In contrast to the DNA-based molecular determination technique of the microbiome, the culture technique depicts active and viable bacteria [20]. In a comprehensive analysis of the microbiota of endodontic infections, the authors of the aforementioned study even showed that some bacterial species, that could be isolated and identified by the culture technique could not be detected by the culture-independent cloning technique. On the other hand, it has been shown that up to 50% of the oral bacteria could not be detected by the culture technique [21]. Zaura et al. [22] showed that healthy individuals exposed to a single antibiotic treatment undergo considerable microbial shifts of the microbiome of their faeces, while their salivary microbiome composition remains more robust and stable. The minor changes in bacterial composition revealed in the present study deliver another hint regarding the stability of the oral microbiota, including in the supragingival oral biofilm.

## 5. Conclusions

Systematically considering the log ratio of concentrations between bacterial pairs has offered new insights into the influence of nutrition on the oral biofilm and the relation of specific bacterial groups with each other. The methods make it possible to circumvent some typical problems encountered in the analysis of compositional data.

**Author Contributions:** A.A.-A., A.A., J.P.W. and E.H. designed the original study. L.K. and A.A. coordinated the sample collection and participated in data analysis. K.V. had the idea for the topic, performed the statistical analysis and designed the figures. All authors edited the manuscript and approved the final article. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the German Research Foundation (DFG, AL-1179/2-1). The article processing charge was funded by the Baden-Württemberg Ministry of Science, Research and Art and the University of Freiburg in the Open Access Publishing funding program. The funding had no influence on the design of the study and collection, analysis and interpretation of data and in writing the manuscript.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the University of Freiburg (Nr. 237/14). All experiments and data collections were performed in accordance with relevant guidelines and regulations.

**Informed Consent Statement:** A written informed consent was obtained from all participants involved in the study.

**Data Availability Statement:** The data are available on request from the authors.

**Acknowledgments:** The authors thank Werner Vach for critical remarks.

**Conflicts of Interest:** The authors report no conflicts of interest.

**Sample Availability:** The data are available on request from the authors.

## Appendix A. Description of the Bacterial Groups

- *Gemella Granulicatella Abiotrophia* spp.: *Gemella morbillorum*, *Gemella haemolysans*, *Gemella sanguinis*, *Granulicatella adiacens*, *Granulicatella elegans*, *Abiotrophia defectiva*
- *Actinomyces* spp.: *Actinomyces oris*, *Actinomyces odontolyticus*, *Actinomyces dentalis*, *Actinomyces georgiae*, *Actinomyces naeslundii*
- *Rothia* spp.: *Rothia mucilaginoso*, *Rothia dentocariosa*, *Rothia aeria*, *Corynebacterium* spp.

- *Neisseria* spp.: *Neisseria macacae/mucosa*, *Neisseria oralis*, *Neisseria subflava*, *Neisseria bacilliformis*, *Neisseria elongata*, *Neisseria flavescens*, *Neisseria perflava*, *Neisseria cinerea*, *Lautrop mirabilis*
- *Capnocytophaga* spp.: *Capnocytophaga granulosa*, *Capnocytophaga gingivalis*, *Capnocytophaga ochracea*, *Capnocytophaga sputigena*
- HACEK: *Haemophilus haemolyticus*, *Haemophilus parahaemolyticus*, *Haemophilus parainfluenzae*, *Haemophilus influenzae*, *Cardiobacterium hominis*, *Eikenella corrodens*, *Kingella* spp.
- *Fusobacterium* spp.: *Fusobacterium nucleatum*, *Fusobacterium periodontium*
- *Campylobacter* spp.: *Campylobacter rectus*, *Campylobacter concisus*, *Campylobacter showae*
- *Streptococcus* spp. group 2: *Streptococcus oralis*, *Streptococcus mitis*, *Streptococcus infantis*, *Streptococcus australis*, *Streptococcus peroris*, *Streptococcus salivarius*, *Streptococcus vestibularis*, *Streptococcus anginosus* group
- *Streptococcus* spp. group 3: *Streptococcus sanguinis*, *Streptococcus parasanguinis*, *Streptococcus gordonii*

## Appendix B. Materials and Methods

### Appendix B.1. Notation

Let  $f_{ip}^b$  denote the fraction of bacterium  $b$  in participant  $i$  and phase  $p$ .

#### Appendix B.1.1. Log Ratio-Based Analysis of Fractions

For each participant  $i$  with  $1, \dots, I$  and phase  $p$ , we consider the log ratios  $l_{ip}^{bc}$  comparing the fractions of the two bacteria  $b$  and  $c$ , i.e.,  $l_{ip}^{bc} = \log_{10} \frac{f_{ip}^b}{f_{ip}^c}$ . The mean  $\mu_p^{bc}$  is estimated by  $\hat{\mu}_p^{bc} = \frac{1}{I} \sum_{i=1}^I l_{ip}^{bc}$  and the standard deviation by  $\hat{\sigma}_p^{bc} = \frac{1}{I-1} \sqrt{\sum_{i=1}^I (l_{ip}^{bc} - \hat{\mu}_p^{bc})^2}$ .

#### Appendix B.1.2. Pairwise Differences in Log Ratio Changes and Their Characteristics

Let us consider the ratio  $\frac{f_{ip}^b}{f_{i1}^b}$  for a bacterium  $b$ , which relates the fractions of phase  $p$  to the baseline phase I. Subsequently, we can build ratios for two different bacteria  $b$  and  $c$ :  $\frac{f_{ip}^b}{f_{i1}^b} / \frac{f_{ip}^c}{f_{i1}^c}$  to compare the changes of these 2 bacteria. Hence,

$$\Delta_{ip}^{bc} = \log_{10} \frac{f_{ip}^b}{f_{i1}^b} - \log_{10} \frac{f_{ip}^c}{f_{i1}^c} \quad (\text{A1})$$

can be interpreted as a difference in log ratio changes:

$$(A) \quad \Delta_{ip}^{bc} = l_{ip}^b - l_{ip}^c \quad \text{with} \quad l_{ip}^b = \log_{10} \frac{f_{ip}^b}{f_{i1}^b}$$

or as changes in the pairwise log ratios

$$(B) \quad \Delta_{ip}^{bc} = l_{ip}^{bc} - l_{i1}^{bc} \quad \text{with} \quad l_{ip}^{bc} = \log_{10} \frac{f_{ip}^b}{f_{ip}^c}$$

We are interested in the mean value and standard deviation of  $\Delta_{ip}^{bc}$ , which can be estimated by  $\hat{\mu}_{\Delta_p}^{bc} = \frac{1}{I} \sum_{i=1}^I \Delta_{ip}^{bc}$  and  $\hat{\sigma}_{\Delta_p}^{bc} = \frac{1}{I-1} \sqrt{\sum_{i=1}^I (\Delta_{ip}^{bc} - \hat{\mu}_{\Delta_p}^{bc})^2}$ . We report t-test-based 95% confidence intervals.



### Appendix B.1.3. Average Log Ratio Changes

The expected pairwise log ratio changes also allow us to characterise the location of each bacterium across all bacteria for each participant by looking at the average, i.e., by considering

$$\mu_{\Delta_p}^b = \frac{1}{B} \sum_{c=1}^B \mu_{\Delta_p}^{bc} \quad (\text{A2})$$

which we call average log ratio changes.

$\mu_{\Delta_p}^b$  (and its estimate  $\hat{\mu}_{\Delta_p}^b$ ) also have by definition the attractive feature, that its mean value over all bacteria  $\bar{\mu}^\bullet$  is equal to 0, because  $\mu^{bc} = -\mu^{cb}$ . This is useful from a pedagogical perspective, as when presenting the estimates of  $\mu^b$  for all bacteria  $b = 1, \dots, B$  together all deviations from 0 sum up to 0, which complies with the behaviour of fractions. For the latter, it is natural that changes in one direction for some bacteria are compensated by changes in the opposite direction for the other bacteria. Bar charts of  $\hat{\mu}^b$  with 95% confidence intervals over the bacteria are used to visualise the distribution of the changes. Similar charts [23,24] have been used to show the distribution of log ratios in cross-sectional analyses.

### Appendix B.2. Evidence of Composition Change

To consider the overall evidence of a change in the composition of the bacteria from baseline to a specific phase, we can consider as a test statistic the sum of the absolute deviations of  $\mu^b$  from 0 over all bacteria, or the sum of the absolute deviations of  $\mu^{bc}$  over all bacterial pairs. The test statistic corresponds to the area of a bar chart type presentations of these values, as we use in this paper.

The null hypothesis of interest corresponds to no effect of the change in nutrition to the bacterial spectrum in each individual. Therefore, the spectra observed at baseline and in the specific phase are simply observations of the same true spectrum blurred by some noise. Consequently, the two observations are exchangeable under the null hypothesis, and the distribution of a test statistic under the null hypothesis can be obtained by allowing randomly exchanging the whole spectrum in each individual. An exact approximation is obtained by considering all  $2^{11}$  possibilities to exchange the spectra in a subset of the individuals. The  $p$ -value can then be determined by counting how often the observed value exceeds the values of the test statistic observed for each possibility.

## References

1. Anderson, A.C.; Rothballer, M.; Altenburger, M.J.; Woelber, J.P.; Karygianni, L.; Vach, K.; Hellwig, E.; Al-Ahmad, A. Long-term fluctuation of oral biofilm microbiota following different dietary phases. *Appl. Environ. Microbiol.* **2020**. [[CrossRef](#)]
2. Bradshaw, D.J.; Lynch, R.J.M. Diet and the microbial aetiology of dental caries: New paradigms. *Int. Dent. J.* **2013**, *63*, 64–72. [[CrossRef](#)]
3. Moynihan, P.; Petersen, P.E. Diet, nutrition and the prevention of dental diseases. *Public Health Nutr.* **2004**, *7*, 201–226. [[CrossRef](#)] [[PubMed](#)]
4. Petti, S.; Simonetti, R.; Simonetti D'Arca, A. The effect of milk and sucrose consumption on caries in 6-to-11-year-old Italian schoolchildren. *Eur. J. Epidemiol.* **1997**, *13*, 659–664. [[CrossRef](#)]
5. Aimutis, W.R. Bioactive Properties of Milk Proteins with Particular Focus on Anticariogenesis. *J. Nutr.* **2004**, *134*, 989S–995S. [[CrossRef](#)] [[PubMed](#)]
6. Reynolds, E.C.; Cai, F.; Shen, P.; Walker, G.D. Retention in Plaque and Remineralization of Enamel Lesions by Various Forms of Calcium in a Mouthrinse or Sugar-free Chewing Gum. *J. Dent. Res.* **2016**. [[CrossRef](#)]
7. Schlafer, S.; Ibsen, C.J.S.; Birkedal, H.; Nyvad, B. Calcium-Phosphate-Osteopontin Particles Reduce Biofilm Formation and pH Drops in situ Grown Dental Biofilms. *Caries Res.* **2017**, *51*, 26–33. [[CrossRef](#)] [[PubMed](#)]
8. Takahashi, N.; Nyvad, B. Caries ecology revisited: Microbial dynamics and the caries process. *Caries Res.* **2008**, *42*, 409–418. [[CrossRef](#)]
9. Marsh, P.D. In Sickness and in Health—What Does the Oral Microbiome Mean to Us? An Ecological Perspective. *Adv. Dent. Res.* **2018**. [[CrossRef](#)] [[PubMed](#)]
10. Mira, A. Oral Microbiome Studies: Potential Diagnostic and Therapeutic Implications. *Adv. Dent. Res.* **2018**. [[CrossRef](#)] [[PubMed](#)]

11. Aas, J.A.; Griffen, A.L.; Dardis, S.R.; Lee, A.M.; Olsen, I.; Dewhirst, F.E.; Leys, E.J.; Paster, B.J. Bacteria of Dental Caries in Primary and Permanent Teeth in Children and Young Adults. *J. Clin. Microbiol.* **2008**, *46*, 1407–1417. [[CrossRef](#)] [[PubMed](#)]
12. Levy, R.M.; Giannobile, W.V.; Feres, M.; Haffajee, A.D.; Smith, C.; Socransky, S.S. The Effect of Apically Repositioned Flap Surgery on Clinical Parameters and the Composition of the Subgingival Microbiota: 12-Month Data. *Int. J. Periodontics Restor. Dent.* **2002**, *22*, 209–219.
13. Koopman, J.E.; van der Kaaij, N.C.W.; Buijs, M.J.; Elyassi, Y.; van der Veen, M.H.; Crielaard, W.; Ten Cate, J.M.; Zaura, E. The Effect of Fixed Orthodontic Appliances and Fluoride Mouthwash on the Oral Microbiome of Adolescents—A Randomized Controlled Clinical Trial. *PLoS ONE* **2015**, *10*, e0137318. [[CrossRef](#)]
14. Velmurugan, S.; Gan, J.M.; Rathod, K.S.; Khambata, R.S.; Ghosh, S.M.; Hartley, A.; Van Eijl, S.; Sagi-Kiss, V.; Chowdhury, T.A.; Curtis, M.; et al. Dietary nitrate improves vascular function in patients with hypercholesterolemia: A randomized, double-blind, placebo-controlled study. *Am. J. Clin. Nutr.* **2016**, *103*, 25–38. [[CrossRef](#)]
15. Greenacre, M.J. *Compositional Data Analysis in Practice*; CRC Press: Boca Raton, FL, USA; Taylor & Francis Group: Abingdon, UK, 2019.
16. Lamont, R.J.; Koo, H.; Hajishengallis, G. The oral microbiota: Dynamic communities and host interactions. *Nat. Reviews. Microbiol.* **2018**, *16*, 745–759. [[CrossRef](#)] [[PubMed](#)]
17. Bernardi, S.; Karygianni, L.; Filippi, A.; Anderson, A.C.; Zürcher, A.; Hellwig, E.; Vach, K.; Macchiarelli, G.; Al-Ahmad, A. Combining culture and culture-independent methods reveals new microbial composition of halitosis patients' tongue biofilm. *MicrobiologyOpen* **2020**, *9*, e958. [[CrossRef](#)] [[PubMed](#)]
18. Anderson, A.C.; Rothballer, M.; Altenburger, M.J.; Woelber, J.P.; Karygianni, L.; Lagkouvardos, I.; Hellwig, E.; Al-Ahmad, A. In-vivo shift of the microbiota in oral biofilm in response to frequent sucrose consumption. *Sci. Rep.* **2018**, *8*. [[CrossRef](#)] [[PubMed](#)]
19. Vach, K.; Al-Ahmad, A.; Anderson, A.; Woelber, J.P.; Karygianni, L.; Wittmer, A.; Hellwig, E. Analysing the Relationship between Nutrition and the Microbial Composition of the Oral Biofilm—Insights from the Analysis of Individual Variability. *Antibiotics* **2020**, *9*, 479. [[CrossRef](#)] [[PubMed](#)]
20. Anderson, A.C.; Hellwig, E.; Vespermann, R.; Wittmer, A.; Schmid, M.; Karygianni, L.; Al-Ahmad, A. Comprehensive Analysis of Secondary Dental Root Canal Infections: A Combination of Culture and Culture-Independent Approaches Reveals New Insights. *PLoS ONE* **2012**, *7*, e49576. [[CrossRef](#)]
21. Paster, B.J.; Boches, S.K.; Galvin, J.L.; Ericson, R.E.; Lau, C.N.; Levanos, V.A.; Sahasrabudhe, A.; Dewhirst, F.E. Bacterial Diversity in Human Subgingival Plaque. *J. Bacteriol.* **2001**, *183*, 3770–3783. [[CrossRef](#)]
22. Zaura, E.; Brandt, B.W.; Teixeira de Mattos, M.J.; Buijs, M.J.; Caspers, M.P.M.; Rashid, M.U.; Weintraub, A.; Nord, C.E.; Savell, A.; Hu, Y.; Coates, A.R.; Hubank, M.; Spratt, D.A.; Wilson, M.; Keijser, B.J.F.; Crielaard, W. Same Exposure but Two Radically Different Responses to Antibiotics: Resilience of the Salivary Microbiome versus Long-Term Microbial Shifts in Feces. *mBio* **2015**, *6*, e01693-15. [[CrossRef](#)] [[PubMed](#)]
23. Chastin, S.; Palarea-Albaladejo, J.; Dontje, M.; Skelton, D. Combined Effects of Time Spent in Physical Activity, Sedentary Behaviors and Sleep on Obesity and Cardio-Metabolic Health Markers: A Novel Compositional Data Analysis Approach. *PLoS ONE* **2015**, *10*, e0139984. [[CrossRef](#)] [[PubMed](#)]
24. Martín Fernández, J.A.; Daunis i Estadella, J.; Mateu i Figueras, G. On the Interpretation of Differences between Groups for Compositional Data. *Sort Stat. Oper. Res. Trans.* **2015**, *39*, 231–252.