Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Food-allergic children harbour a reduced oral bacterial load altering taxa frequencies in 16S rRNA amplicon sequencing

Clara Bonaiti
University of Milan
Simona Panelli
University of Milan
Lodovico Sterzi
University of Milan
Miriam Acunzo
Ospedale dei Bambini Vittore Buzzi
Enza D'Auria
Ospedale dei Bambini Vittore Buzzi
Gianvincenzo Zuccotti
Ospedale dei Bambini Vittore Buzzi
Francesco Comandatore
francesco.comandatore@unimi.it
University of Milan

Article

Keywords: food allergies, oral microbiota, 16S rRNA amplicon sequencing, relative frequencies, absolute abundances, total bacterial load

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11349955/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Food-allergic children harbour a reduced oral bacterial load altering taxa frequencies in 16S rRNA amplicon sequencing

Clara Bonaiti[1], Simona Panelli[1], Lodovico Sterzi[1], Miriam Acunzo1[1,2], Enza D'Auria[2], Gianvincenzo Zuccotti[1,2], Francesco Comandatore*[1]

[1]Department of Biomedical and Clinical Sciences, Pediatric Clinical Research Center "Romeo and Enrica Invernizzi", University of Milan, 20157, Milan, Italy

[2]Department of Pediatrics, Buzzi Children's Hospital, University of Milan, 20154, Milan, Italy.

Correspondence to: Prof. Francesco Comandatore, Department of Biomedical and Clinical Sciences, Pediatric Clinical Research Center "Romeo and Enrica Invernizzi", University of Milan, 20157, Milan, Italy. Email: francesco.comandatore@unimi.it

Abstract

The microbiota is increasingly recognized as a key factor in food allergy. However, the use of taxon frequencies to describe microbiota composition can lead to autocorrelated values, potentially misleading results and associations. Indeed, this approach assumes that the total bacterial load is relatively constant across the groups being compared, an assumption that may not hold in some pathologies. For this reason, we extended our previous work comparing the oral microbiota of food-allergic and control children by estimating the absolute abundances of bacterial taxa rather than their relative frequencies. Absolute abundances were estimated by combining pan-bacterial 16S rRNA gene qPCR with amplicon sequencing. We found that food-

allergic children exhibited a marked reduction in total oral bacterial load, resulting in a picture that only partially overlapped with our previous findings. Our data confirm that variations in bacterial load across samples hinder efforts to associate microbiota features with disease phenotypes. We also identify reduced oral bacterial load as a previously unrecognized characteristic of children with food allergies. This aligns with reports in other immune-mediated disorders, suggesting that overall microbiota depletion may represent a defining, potentially diagnostic, feature of disease-associated ecosystems.

Keywords: food allergies, oral microbiota, 16S rRNA amplicon sequencing, relative frequencies, absolute abundances, total bacterial load

Introduction

Food allergies (FA) are inappropriate immune reactions triggered by the ingestion of normally innocuous food antigens, resulting in local or systemic reactions. Symptoms range from life threatening conditions, such as anaphylaxis, to respiratory, cutaneous and gastrointestinal disorders [1]. FA prevalence can reach up to 10%, with considerable variation across geographical regions, allergens and age groups [2]. Over the past century, FA prevalence has significantly increased in western countries [3] emerging as a major public health concern.

FA are complex conditions resulting from the interplay of multiple endogenous and exogenous influences that can either attenuate or exacerbate their pathogenesis [4,5]. Early-life factors, pollution, chemicals, diet, infections and interactions with other organisms have all been implicated in the pathogenesis [6]. Microbial exposure is now recognized as a critical environmental risk modifier, with both pathogens and commensal bacteria generally endowed with protective roles [7,8]. Recent works point to a central role for the microbiota [8], through multiple mechanisms. For example the microbiota exerts immunomodulatory effects *via* the

production of metabolites such as short chain fatty acids (SCFA), which promote tolerance and possess anti-inflammatory activities[9]. A dysregulated homeostatic interaction between the host and its microbiota has been identified as a predisposing factor for FA, with microbial dysbiosis playing a critical role in the establishment, or failure, of oral tolerance [8].

Recent studies have highlighted the pivotal role of the oral host-commensal environment [10]. Using amplicon sequencing on the V3-V4 regions of the 16S rRNA gene, our group reported alterations in the oral bacterial profile of children with FA compared to healthy controls [11]. The most statistically significant difference was the marked enrichment, in allergic children, of "*Candidatus* Gracilibacteria", belonging to the Candidate Phyla Radiation (CPR, also known as *Patescibacteria*), a peculiar and still poorly characterized bacterial clade [12]. A second *Patescibacteria* lineage, "*Candidatus* Saccharimonadia" (formerly TM7), exhibited a trend to increase in the allergic group. Members of *Patescibacteria*, in particular "*Ca.* Saccharimonadia", are increasingly associated with inflammatory diseases, and growing evidence points to their involvement in immunomodulatory processes [13]. Despite this, *Patescibacteria* remain largely elusive and understudied for several methodological and biological reasons [12]. These considerations prompted us to further investigate their enrichment within the oral microbiota of allergic children.

Another key premise for the present work came from a paper by Bruijning and colleagues (2023) [14] which highlighted a potential source of bias in the way microbiota data are typically presented in 16S rRNA amplicon sequencing studies. The 16S rRNA amplicon sequencing can not provide absolute abundances of bacterial taxa, but it expresses them as percentages. The authors evidenced that using relative abundances can bias the comparison of taxa abundance and of their changing, because percentages are inherently autocorrelated [15]. Because the sum of relative abundances in a sample must equal one, a decrease in the relative abundance of one taxon necessarily results in an apparent increase in another, regardless of their actual quantities. This can introduce distortions when assessing correlation between variables, or identifying taxonomic signatures [16,17]. Bruijning and colleagues (2023) [14] propose the use of absolute, rather than relative, abundances, to alleviate these distortions. One way they propose is

integrating microbial relative abundance data with estimates of the total bacterial load of the sample derived from quantitative PCR on the 16S rRNA gene.

As studies comparing absolute vs. relative abundances have begun to emerge [18,19], we extended our investigation on the oral microbiota in food allergy [11]. Previous works, e.g. Marotz et al. (2021) [20] and Vandeputte et al. (2017) [21], have shown that qPCR targeting the 16S rRNA gene provides a reliable proxy for estimating bacterial cell numbers. Accordingly, and also following the approach proposed by Bruijning and colleagues (2023)[14], we used this strategy to estimate the total bacterial load in each sample, and integrated these measurements with the corresponding 16S rRNA V3–V4 amplicon sequencing data (see also [22]).

Determination of the total bacterial loads revealed a reduction in oral bacterial burden in children with FA that was not apparent from relative abundance data alone. Incorporating absolute abundance estimates also partially reshaped the taxonomic profiles obtained from relative abundance data [23], highlighting the influence of overall bacterial load on the interpretation of microbiota composition. These findings underscore the importance of considering absolute abundances when characterizing the microbiota and suggest that a reduced oral bacterial burden may be a feature of FA. More broadly, our results provide a different perspective on the relationship between the host and the oral microbiota relationship in the context of FA.

Results

Estimation of total bacterial load in salivary DNA from food-allergic and control children

The total bacterial load of the 59 salivary DNA samples, collected from children with FA and controls, previously characterized by D'Auria et al. [11], was estimated by qPCR using pan-bacterial primers on the 16S rRNA gene, as detailed in the Methods section.

The bacterial load inferred after the corrections (see Methods) resulted significantly correlated to the 16S rRNA total quantification obtained by the qPCR assay (Supplementary Fig. S1), with a slope of 3.7 and an intercept of $11x10^4$.

As shown in Fig. 1, children with FA displayed a significantly lower oral bacterial load than controls (Wilcoxon–Mann–Whitney test, p-value < $2.2x10^{-16}$). In particular, the median bacterial quantification for 2-µl samples was $7.4x10^{6}$ for allergic children and $9.3x10^{6}$ for the control group, corresponding respectively to $3.7x10^{8}$/ml and $4.7x10^{8}$/ml.

Comparison between the ecological indexes of β-diversity computed from relative and absolute bacterial taxa abundances in food-allergic and control children

The Principal Coordinates Analysis (PCoA), computed both on relative and absolute abundances at the family and genus taxonomic levels, is reported in Fig. 2. For all the four PCoAs, the PERMANOVA test coherently reported the statistically significant separation between the oral microbiota of allergic and control children.

Comparison between relative and absolute bacterial taxa abundances in allergic and control individuals

For each of the 59 saliva samples, the absolute abundance of each bacterial taxon was estimated combining the total bacterial load and the relative abundances of bacterial taxa retrieved from [11] (see Methods for details).

The 16S rRNA amplicon sequencing analysis identified a total of 303 bacterial taxa across all taxonomic levels. Among these, 192 taxa (63%) showed consistent patterns using relative and absolute abundances (Supplementary Table S1). For example, the phylum *Firmicutes* was significantly less abundant in allergic children than in controls, according to both relative and absolute abundances. Overall, coherent results for relative and absolute abundances were observed for eight of 11 phyla (73%), 15/20 classes (75%), 31/47 orders (66%), 45/70 families (64%) and 93/147 genera (63%) (Supplementary Table S1).

Among the 303 taxa identified, 144 showed significant differences in relative and/or absolute abundance between allergic and control children. These comprised four phyla, 8 classes, 22 orders, 43 families, and 67 genera. Coherent differences between the two groups, detected by

both absolute and relative measures, were observed for one of the four phyla (17%), three of the eight classes (36%), 10 of the 22 orders (45%), 20 of the 43 families (47%), and 36 of the 67 genera (54%) (Supplementary Table S1).

At the phylum level, among the six taxa mentioned above, only *Firmicutes*, *Actinobacteriota* and *Bacteroidota* showed concordant results. *Firmicutes* exhibited lower relative and absolute abundances in allergic children, whereas the other phyla displayed comparable values between the two groups (Fig. 3). *Spirochetota* were reduced in allergic children only when considering absolute abundances. In contrast, *Patescibacteria* and unclassified Bacteria displayed higher relative abundances in allergic children, while their absolute abundances remained unchanged. *Firmicutes* was the only phylum that resulted significantly different in allergic children vs controls according to both absolute and relative abundances. As reported in Supplementary Table S1, a total of 98 *Firmicutes* taxa were identified. Of these, 52 exhibited no significant differences between allergic and control children in either relative and absolute abundances, while 15 taxa yielded discordant results. The remaining 31 *Firmicutes* taxa presented significant between-group differences in both analyses. Of these 31 taxa, 27 were significantly reduced in allergic children, mirroring the overall phylum-level trend. These included, for example, the orders *Lactobacillales*, *Staphylococcales*, and *Lachnospirales*; the families *Aerococcaceae*, *Carnobacteriaceae*, *Gemellaceae*, and *Lachnospiraceae*; and the genera *Solobacterium*, *Granulicatella*, *Streptococcus*, and *Gemella*. The remaining four taxa were low-abundance, unclassified *Firmicutes* lineages, and were the only ones resulting significantly enriched in allergic children according to both relative and absolute abundances.

General patterns of concordance/discordance between absolute and relative abundances are shown in the heatmap of Fig. 4 for all taxonomic levels. *Patescibacteria* and sequences unresolved by the 16S rRNA V3-V4 amplicon sequencing represented the portion of the oral microbiota stable in terms of absolute abundance in allergic children, despite the marked, overall reduction of the oral bacterial load in this group. The taxonomic identity of these unresolved sequences was explored as described in Materials and Methods. Fig. 5 shows the distribution of the phyla corresponding to their closest BLAST matches. As evident from Fig. 5, these sequences represent a heterogeneous set of bacterial lineages spanning a broad range of the

known bacterial diversity. The most represented phylum is *Firmicutes*, for both allergic and control children, but several other phyla are also detected.

Discussion

The oral microbiota, a key component of the oral ecosystem, is emerging as a crucial factor associated with the pathogenesis and progression of FA [24]. In our previous work [23], we reported significant differences in the salivary microbiota of children with FA compared with healthy controls, using the relative taxa abundances generated by standard 16S V3-V4 rRNA amplicon sequencing [11]. However, increasing evidence has revealed important limitations of relative abundance-based analyses, that represent the typical way microbiota data are presented in 16S rRNA amplicon sequencing studies. These analyses produce compositional data that can introduce distortions, lead to biased interpretations, and hinder a comprehensive understanding of host-microbiota interactions [14]. Therefore, to expand our investigation on the role of oral microbiota in FA, we determined total bacterial loads and the absolute abundances of individual bacterial taxa in the salivary DNA samples previously characterized by 16S rRNA V3-V4 amplicon sequencing [11].

To this end,we used a qPCR-based approach to quantify the 16S rRNA gene copies. Previous studies have established that this approach provides a reliable proxy for estimating total bacterial loads, and have applied it for different purposes. For example for monitoring changes in the bacterial load of human saliva over time and following acute perturbations [20]; for integrating absolute bacterial measurements with conventional 16S rRNA-based relative abundance data to assess the effects of sample storage conditions [18,19]; for comparing fecal microbial load between healthy individuals and patients with Crohn's disease [21]. In parallel, other studies have demonstrated the stability and suitability of DNA samples for qPCR analyses following long-term storage, e.g. preservation at -20°C for more than 10 years [25,26]. However, a known limitation of the panbacterial 16S rRNA qPCR assay is the variability in 16S rRNA gene copy number among bacterial taxa, which can bias abundance estimates. To overcome this challenge, similarly to what done by Maghini and colleagues [19], we combined panbacterial 16S rRNA

qPCR quantification of the total bacterial load, 16S rRNA amplicon sequencing–determined taxa frequencies, and the taxon-specific 16S rRNA gene copy numbers.

We found that allergic children harbor a significantly lower salivary bacterial load than healthy controls, a key finding that could not have been detected through conventional 16S rRNA amplicon sequencing analyses. The reduced bacterial load may reflect the altered immunological environment in FA, including a T helper 2 (Th2)–skewed immune profile and impaired mucosal homeostasis. Several allergy-associated alterations could influence oral microbial colonization, including altered epithelial barrier function, salivary antimicrobial peptide profiles, and microbial metabolite production (e.g. SCFAs) [24,27]. Furthermore, the low-grade mucosal inflammation associated with FA may alter the nutritional and physicochemical properties of saliva, contributing to the reduced bacterial load [24].
Reduced total bacterial loads have been reported in other immune-related conditions, e.g. in the gut microbiota of patients with Crohn's disease [21] and in the salivary microbiota of children with elevated salivary glucose levels, at risk for Type 2 Diabetes [28]. These studies emphasize the importance of quantifying microbial loads when assessing microbiota alterations and suggest that changes in overall bacterial abundance may represent a hallmark of disease-associated microbiota configurations. This raises the possibility of using total- or taxon-specific abundances for diagnosis or prognostic purposes [28]. Notably, Goodson and colleagues [28] found that total salivary bacterial load predicted elevated oral glucose levels and obesity-related metabolic syndrome more accurately than conventional clinical measures [28,29].

Absolute abundances of bacterial taxa in samples were then evaluated by combining the total bacterial load with taxa relative abundances from our previous work, by following the flow described in [14]. Taxa absolute abundances were used to re-evaluate β-diversity and taxonomic profiles.

Regarding β-diversity, the PERMANOVA on the PCoA continued to show a significant separation between allergic and control children when using the absolute abundances.

Regarding taxonomic profiles, the use of absolute abundances partially revised our previous findings [11]. Approximately 40% of the taxa displayed discordant patterns between allergic and control children when relative and absolute abundances were compared. For example, relative abundances suggested that *Patescibacteria* and unclassified lineages were enriched in allergic children. However, absolute quantification revealed that this apparent increase was an artefact of the reduced overall bacterial load, as the absolute abundances of these taxa remained essentially unchanged between the two groups. For the same reason, reductions in *Actinobacteriota*, *Bacteroidota* and *Spirochetota* in allergic children were evident only when absolute abundances were considered, and were obscured in relative-frequency comparisons. *Firmicutes* represented a notable exception, as their abundance was significantly reduced in allergic children regardless of the quantification method, suggesting a decrease that exceeded the overall reduction in bacterial load.

The differential contribution of individual phyla to the overall reduction of salivary bacterial loads raises several noteworthy points. Most of the major oral bacterial phyla were affected by this decrease, which reached its maximum for *Firmicutes*. This observation is consistent with current knowledge of the immunological environment associated with allergic diseases and the role of *Firmicutes* in maintaining mucosal homeostasis[24,27]. We also observed a marked depletion of *Firmicutes* taxa with known anti-inflammatory properties, including SCFA-producing families (*Lachnospiraceae* and *Ruminococcaceae).*

Patescibacteria and unclassified bacterial lineages remain stable in absolute abundance in the allergic group, despite the marked reduction in total bacterial load. The apparent resilience of these taxa to the perturbations that affect several major oral bacterial phyla is intriguing. This stability may reflect occupation of specific ecological niches, distinctive metabolic or lifestyle traits that allow them to better adapt to the new conditions, or particular interactions with the host immune system [27].

Concerning *Patescibacteria*, our findings are consistent with reports of increased relative abundances of *"Ca.* Saccharimonadia" in dysbiotic and inflammatory microbiomes [12]. "*Ca*.

Saccharimonadia" have been described as "inflammophilic bacteria" capable of thriving in inflammation-associated, nutrient-rich niches [13][30,31]. They also exert important immunomodulatory effects, including the regulation of cytokine responses and the immunoactivation of oral epithelial cells [30,31]. Furthermore, recent evidence shows that "*Ca*. Saccharimonadia" can be internalized by oral epithelial cells and survive inside them [31], a mechanism that may contribute to their stable absolute abundances.

Concerning unclassified bacterial lineages, their taxonomic characterization revealed a heterogeneous origin, spanning all major oral bacterial phyla [27]. This may suggest that factors beyond simple taxonomic affiliation, such as niche specialization or interactions with the host immune system, contribute to their persistence.

In conclusion, the present work builds on recent evidence showing that the use of relative abundances in 16S rRNA sequencing-based microbiome studies can bias comparisons of bacterial taxa and underscores the importance of incorporating absolute quantifications [19]. Despite this, most microbiome studies still rely exclusively on relative taxa abundances [18,19,32]. We investigated the salivary microbiota of allergic children considering these issues, and developing an analytical procedure to estimate total bacterial loads in DNA samples subjected to 16S rRNA sequencing. The method combines qPCR-based quantification of total 16S rRNA gene copies, genus-specific 16S rRNA gene copy numbers retrieved from public databases, and genus-level taxonomic profiles obtained through 16S rRNA V3–V4 amplicon sequencing. We confirmed that substantial variation of bacterial loads can hinder efforts to associate microbiota features with the phenotype under investigation by forcing spurious statistical trends. Importantly, this approach enabled the identification of a key biological characteristic of FA children, undetected in previous analyses based on relative abundances [11]. Specifically, FA children exhibited significantly lower total oral bacterial loads than healthy controls. This observation is consistent with previous reports in Crohn's disease [21] and Type 2 Diabetes [28], and it adds to the list a disease that is experiencing a dramatic increase in prevalence. More broadly, our results support the notion that alterations in overall microbiota

abundance may represent a distinctive feature of disease-associated microbial ecosystems, with potential diagnostic and predictive value [21,28,29].

Methods

Ethics statement

The research was conducted according to the declaration of Helsinki, and all methods followed the relevant guidelines and regulations. The Ethics Committee of ASST-Fatebenefratelli-Sacco approved the study (Ref. n. 2021/ST/041). Privacy rights of subjects were carefully observed and authors did not have access to information that could identify individual participants. During the recruitment of subjects, parents signed a written informed consent to allow their child to participate in the study. Samples and data were first accessed for research purposes on 01 February 2022.

Dataset

The study included the same 59 salivary DNA samples previously characterized by D'Auria et al. [11]. In that study, DNA was extracted from saliva collected from 29 children with FA and 30 healthy controls, and analysed by 16S rRNA V3-V4 amplicon sequencing (data available under the Project Accession ID PRJEB128586). More details and demographic data of the cohort are available in [11]. To ensure direct comparability between the qPCR results and the amplicon sequencing analysis, the original DNA tubes were used for the present study. Samples were not re-extracted in order to avoid any kind of technical variability.

Total bacterial load estimation

For each sample, the total bacterial load was estimated by integrating (i) the total number of 16S rRNA gene copies measured by qPCR, (ii) genus-specific 16S rRNA gene copy numbers retrieved from public databases, and (iii) genus-level taxonomic profiles obtained through 16S rRNA V3–V4 amplicon sequencing.

The total number of 16S rRNA gene copies was measured *via* qPCR using the 926F/1062R pan-bacterial primers [33][34]. Specifically, the total number of 16S rRNA gene copies were determined in 2-µL DNA samples, in a qPCR volume of 15 µL, also containing 7.5 µL of 2x SsoAdvanced Universal SYBR® Green Supermix (BioRad, Hercules, CA), 4.7 µL of PCR amplification-grade water (Promega Corporation, Wisconsin, USA), and 0.4 µL of 10 µM stocks for the pan-bacterial primers 926F and 1062R [33]. Each sample was qPCR-amplified in three technical replicates. The qPCR assays were performed on a BioRad CFX Connect real-time PCR System (BioRad, Hercules, CA) with the following profile: initial denaturation (95°C 5 min), 40 cycles at 95°C 15s, 61.5°C 15s, 72°C 20s. Finally, a melting curve was generated, with a range of temperature between 60° and 95°C.

Because of the time interval between V3–V4 amplicon sequencing and the qPCR experiments, which was approximately one year, we assessed the stability and suitability of the DNA samples for qPCR analysis following storage at −20 °C. The qPCR assays were therefore repeated in September 2026, immediately before submission of the present manuscript, approximately four years after the original qPCR analyses. Repeat measurements were performed on the 15 DNA samples that were still available (5 children with FA and 10 healthy controls). As shown in Supplementary Fig. S2, the bacterial load estimates obtained in the two qPCR assays were significantly correlated, consistent with previous reports [25,26].

Although the total number of 16S rRNA gene copies can provide a raw indication of the number of bacteria in a sample, this evaluation may be unprecise because different bacterial species can harbor different numbers of 16S rRNA gene copies (from 1 to >10). To account for this variability, the number of 16S rRNA copies for each genus was retrieved from the rrnDB 5.10 database [35].

Lastly, genus-level relative abundances were derived from the 16S rRNA V3–V4 amplicon sequencing dataset reported by D'Auria et al.[11]. Sequence reads were processed using the Mothur pipeline and clustered into Operational Taxonomic Units (OTUs) prior to taxonomic assignment. This approach is aimed to reduce the impact of variability in 16S rRNA gene copy numbers across taxa, improving the accuracy of genus-level abundance estimates.

Thus, these OTU frequencies were used to estimate the total bacterial load as follows

337 (1) $$N_{OTU} = \frac{N_{qPCR}}{\sum_i^g f_{OTUi} * n_i}$$

338

339 With N_{OTU} = total bacterial load computed using OTU counts, N_{qPCR} = total number of 16S rRNA
340 copies obtained from qPCR; $f_{OTU\,i}$ = frequency of the genus i computed using OTUs; n_i = median
341 number of 16S rRNA copies present into the genus i; g = total number of genera.

342 The Mothur pipeline also returns, as an intermediate file, the quantification of the 16S rRNA reads
343 assigned to each genus. Even if these quantifications are particularly affected by 16S rRNA gene
344 copies variation among genera, it is possible to correct this distortion using the genera 16S rRNA
345 copies obtained above.

346 Thus, the 16S rRNA reads percentages were used to the total bacterial load as follows:

347 (2) $$N_{reads} = N_{qPCR} * \sum_i^g \frac{P_{Ri}}{n_i}$$

348 With N_{reads} = total bacterial load computed using reads counts; $P_{R\,i}$ = percentage of 16S rRNA
349 reads assigned to the genus i.

350 As the two total bacterial loads resulted highly congruent (Pearson correlation, p value < 0.01,
351 Supplementary Fig. S1), the total bacterial load was computed as the average of the two
352 estimations, as follows

353 (3) $$N = \frac{(N_{OTU} + N_{reads})}{2}$$

354

355 All commands and files are available at
356 https://github.com/FrancescoComandatore/Bacterial_absolute_quantification.

357

358 **Beta diversity analysis**

359 The total bacterial loads obtained for allergic and control children were compared using the
360 Wilcoxon–Mann–Whitney test, with a significance threshold (p-value) set to 0.05. The absolute
361 abundance of each taxon in each sample was then estimated by multiplying the total bacterial

load (obtained as described above) by the relative frequencies obtained from the 16S V3-V4 amplicon sequencing retrieved from D'Auria et al. 2023[11].

Taxonomic pictures and ecological indices of β-diversity obtained using both absolute and relative abundances were computed using the R library Vegan (https://CRAN.R-project.org/package=vegan). The α-diversity indices were not recalculated as they are not expected to change when considering absolute abundances, obtained by multiplying relative frequencies by a constant. Concerning the microbial profiles obtained by relative and absolute abundances, they were compared using Mann–Whitney U test, with a significance threshold (p-value) set to 0.05.

OTU sequences classified as Unclassified at phylum level were retrieved and BLASTn searched against the SILVA v138.1 database [36] to retrieve the closest phylum.

References

1. Sicherer, S. H. & Sampson, H. A. Food allergy: Epidemiology, pathogenesis, diagnosis, and treatment. *J Allergy Clin Immunol* **133**, 291–307; quiz 308 (2014).
2. Sicherer, S. H. Epidemiology of food allergy. *J Allergy Clin Immunol* **127**, 594–602 (2011).
3. Prescott, S. & Allen, K. J. Food allergy: riding the second wave of the allergy epidemic. *Pediatr Allergy Immunol* **22**, 155–160 (2011).
4. Holloway, J. W., Yang, I. A. & Holgate, S. T. Genetics of allergic disease. *J Allergy Clin Immunol* **125**, S81–94 (2010).
5. Falcon, R. M. G. & Caoili, S. E. C. Immunologic, genetic, and ecological interplay of factors involved in allergic diseases. *Front Allergy* **4**, 1215616 (2023).
6. Peters, R. L., Mavoa, S. & Koplin, J. J. An Overview of Environmental Risk Factors for Food Allergy. *Int J Environ Res Public Health* **19**, (2022).
7. Azad, M. B. *et al.* Infant gut microbiota and food sensitization: associations in the first year of life. *Clin Exp Allergy* **45**, 632–643 (2015).
8. Shu, S.-A. *et al.* Microbiota and Food Allergy. *Clin Rev Allergy Immunol* **57**, 83–97 (2019).

9. Tan, J. *et al.* Dietary Fiber and Bacterial SCFA Enhance Oral Tolerance and Protect against Food Allergy through Diverse Cellular Pathways. *Cell Rep* **15**, 2809–2824 (2016).

10. Ho, H.-E. *et al.* Multidimensional study of the oral microbiome, metabolite, and immunologic environment in peanut allergy. *J Allergy Clin Immunol* **148**, 627–632.e3 (2021).

11. D'Auria, E. *et al.* Alteration of taste perception, food neophobia and oral microbiota composition in children with food allergy. *Sci Rep* **13**, 7010 (2023).

12. Naud, S. *et al.* Candidate Phyla Radiation, an Underappreciated Division of the Human Microbiome, and Its Impact on Health and Disease. *Clin Microbiol Rev* **35**, e0014021 (2022).

13. Chipashvili, O. *et al.* Episymbiotic Saccharibacteria suppresses gingival inflammation and bone loss in mice through host bacterial modulation. *Cell Host Microbe* **29**, 1649–1662.e7 (2021).

14. Bruijning, M. *et al.* Relative abundance data can misrepresent heritability of the microbiome. *Microbiome* **11**, 222 (2023).

15. 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).

16. Carr, A., Diener, C., Baliga, N. S. & Gibbons, S. M. Use and abuse of correlation analyses in microbial ecology. *ISME J* **13**, 2647–2655 (2019).

17. Nearing, J. T. *et al.* Microbiome differential abundance methods produce different results across 38 datasets. *Nat Commun* **13**, 342 (2022).

18. Rao, C. *et al.* Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature* **591**, 633–638 (2021).

19. Maghini, D. G. *et al.* Quantifying bias introduced by sample collection in relative and absolute microbiome measurements. *Nat Biotechnol* **42**, 328–338 (2024).

20. Marotz, C. *et al.* Quantifying Live Microbial Load in Human Saliva Samples over Time Reveals Stable Composition and Dynamic Load. *mSystems* **6**, (2021).

21. Vandeputte, D. *et al.* Quantitative microbiome profiling links gut community variation to microbial load. *Nature* **551**, 507–511 (2017).

22. Maghini, D. G. *et al.* Author Correction: Quantifying bias introduced by sample collection in relative and absolute microbiome measurements. *Nat Biotechnol* **42**, 1313 (2024).

23. D'Auria, E. *et al.* Alteration of taste perception, food neophobia and oral microbiota composition in children with food allergy. *Sci Rep* **13**, 7010 (2023).

24. Zhang, R. *et al.* The role of the oral microbiota in allergic diseases: Current understandings and future trends. *Microbiol Res* **299**, 128254 (2025).

25. Sguazzi, G. *et al.* Microbial DNA in human nucleic acid extracts: Recoverability of the microbiome in DNA extracts stored frozen long-term and its potential and ethical implications for forensic investigation. *Forensic Sci Int Genet* **59**, 102686 (2022).

26. Sguazzi, G. *et al.* Can DNA withstand the test of time? Exploring degradation across storage conditions. *Forensic Sci Int Genet* **84**, 103520 (2026).

27. Baker, J. L., Mark Welch, J. L., Kauffman, K. M., McLean, J. S. & He, X. The oral microbiome: diversity, biogeography and human health. *Nat Rev Microbiol* **22**, 89–104 (2024).

28. Goodson, J. M. *et al.* The salivary microbiome is altered in the presence of a high salivary glucose concentration. *PLoS One* **12**, e0170437 (2017).

29. Goodson, J. M. *et al.* Metabolic disease risk in children by salivary biomarker analysis. *PLoS One* **9**, e98799 (2014).

30. He, X. *et al.* Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. *Proc Natl Acad Sci U S A* **112**, 244–249 (2015).

31. Chouhan, D. *et al.* Episymbiotic Saccharibacteria suppresses epithelial immunoactivation through Type IV pili and TLR2 dependent endocytosis. *bioRxiv* (2025) doi:10.1101/2025.05.30.656655.

32. Vandeputte, D. *et al.* Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. *Nat Commun* **12**, 6740 (2021).

33. Bacchetti De Gregoris, T., Aldred, N., Clare, A. S. & Burgess, J. G. Improvement of phylum- and class-specific primers for real-time PCR quantification of bacterial taxa. *J Microbiol Methods* **86**, 351–356 (2011).

34. Papaleo, S. *et al.* Comparison of qPCR protocols for quantification of 'Candidatus Saccharibacteria', belonging to the Candidate Phyla Radiation, suggests that 23S rRNA is a better target than 16S rRNA. *PLoS One* **19**, e0310675 (2024).

35. Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K. & Schmidt, T. M. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res* **43**, D593–8 (2015).

36. Beccati, A., Gerken, J., Quast, C., Yilmaz, P. & Glöckner, F. O. SILVA tree viewer: interactive web browsing of the SILVA phylogenetic guide trees. *BMC Bioinformatics* **18**, 433 (2017).

Acknowledgments

We wish to thank the Romeo ed Enrica Invernizzi Foundation for support.

Author Contributions

CB: Investigation, Writing – original draft, Writing – review & editing; SP: Conceptualization, Writing – review & editing; LS: Data curation; MA: Investigation; ECD: Conceptualization, Investigation; GZ: Conceptualization, Supervision; FC: Conceptualization, Data curation, Supervision, Writing – original draft, Writing – review & editing.

All authors approved the final version of the manuscript.

Data Availability Statement

All R commands and files used for the estimation of absolute abundances are available at https://eur02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fgithub.com%2FFrancescoComandatore%2FBacterial_absolute_quantification&data=05%7C02%7Cfrancesco.comandatore%40unimi.it%7Cd21319be9c3e4335901f08df2ff830b1%7C13b55eef70184674a3d7cc0db06d545c%7C0%7C0%7C639282415162094462%7CUnknown%7CTWFpbGZsb3d8eyJFbXB0eU1hcGkiOnRydWUsIlYiOiIwLjAuMDAwMCIsIlAiOiJXaW4zMiIsIkFOIjoiTWFpbCIsIldUIjoyfQ%3D%3D%7C0%7C%7C%7C&sdata=6WkiREpXO68%2FMInnohbmNpWgnLCKV2WXV4V172laAK4%3D&reserved=0.

The sequences are currently available on the ENA website PRJEB128586 (they will be available on NCBI soon). You can find the project ID page at this link: https://eur02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.ebi.ac.uk%2Fena%2Fbrowser%2Fview%2FPRJEB128586&data=05%7C02%7Cfrancesco.comandatore%40unimi.it%7Cd21319be9c3e4335901f08df2ff830b1%7C13b55eef70184674a3d7cc0db06d545c%7C0%7C0%7C639282415162114179%7CUnknown%7CTWFpbGZsb3d8eyJFbXB0eU1hcGkiOnRydWUsIlYiOiIwLjAuMDAwMCIsIlAiOiJXaW4zMiIsIkFOIjoiTWFpbCIsIldUIjoyfQ%3D%3D%7C0%7C%7C%7C&sdata=B7pO0FWRIYwXgLNxmimWQTcbN6jBjLrj%2FzvrHMmrUxE%3D&reserved=0

Additional information

The oline version contains supplementary material

Funding

The study has been financially supported by the University of Milan by the project Grandi Sfide di Ateneo (GSA)-Linea 6.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Figure Legends

Fig. 1. Absolute oral bacterial abundances in children with food allergy and healthy controls.

Log10 of absolute bacterial abundances estimated as the 16S rRNA gene copies per 2-µl samples determined by pan-bacterial qPCR. Data for the two groups were compared using the Wilcoxon–Mann–Whitney test.

Fig. 2. β-diversity analysis.

The microbiota composition of allergic vs control children were compared by Principal Coordinates Analysis (PCoA). (a-b) PCoA at the taxonomic level of the genus; (c-d) PCoA at the taxonomic level of the family. PCoA obtained for relative abundances are shown in plots a and c (samples from allergic children are colored in dark blue and those from controls in light blue),

while those computed from absolute abundances are shown in plots b and d (allergic in dark green and controls in light green). Allergic vs control microbiota compositions were compared using the PERMANOVA test and significant threshold was set < 0.05.

Fig. 3. Absolute versus relative abundances of bacterial phyla

Plots a-f show relative and absolute abundances of bacterial phyla with significant differences in relative and/or absolute abundances between control and allergic children. In each plot, two couples of boxplots are shown: on the left, boxplots comparing the relative abundances between allergic (dark blue) and control (light blue) subjects; on the right, the boxplots comparing the absolute abundances between allergic (dark green) and control (light green) subjects.

Fig. 4. Concordance between absolute and relative abundances

Heatmap showing patterns of concordance and discordance between relative and absolute abundances across all taxonomic levels. Rows represent relative abundance comparisons, and columns represent absolute abundance comparisons. Row and column labels refer to comparisons between allergic (A) and control (C) children: A > C indicates taxa with statistically significantly higher abundance in allergic individuals; A = C indicates no statistically significant difference; A < C indicates taxa with significantly lower abundance in allergic individuals compared with control individuals. Values within cells indicate the number of taxa exhibiting the pattern defined by each row–column combination.

Fig. 5. Taxonomic assignment of unclassified bacteria

Distribution of the bacterial phyla closest to sequences unclassified at the phylum level following 16S rRNA V3-V4 amplicon sequencing. The allergic group is shown on the left, and the control group on the right.

Figures

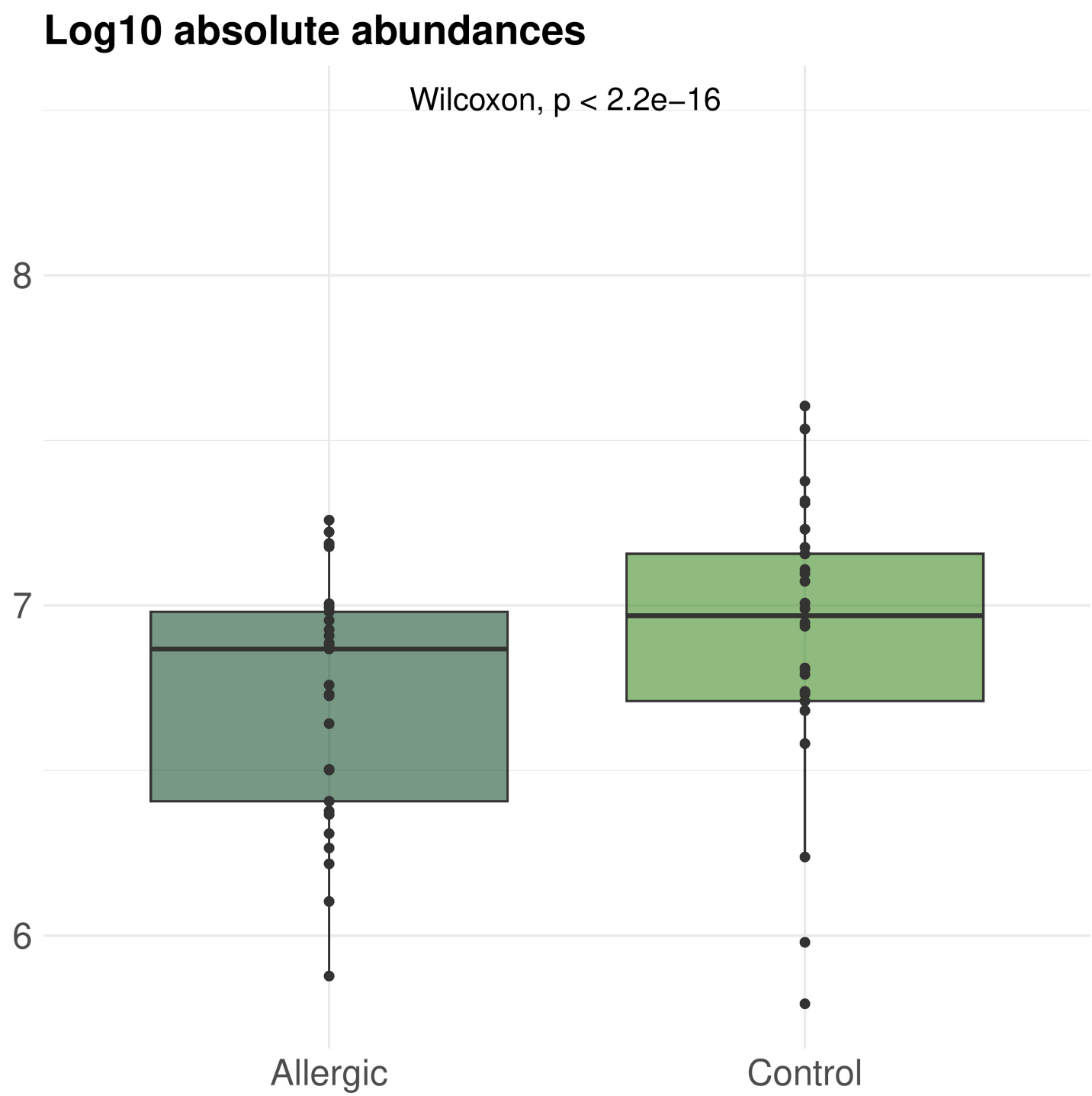


Figure 1

Absolute oral bacterial abundances in children with food allergy and healthy controls.

Log10 of absolute bacterial abundances estimated as the 16S rRNA gene copies per 2-µl samples determined by pan-bacterial qPCR. Data for the two groups were compared using the Wilcoxon–Mann–

Whitney test.

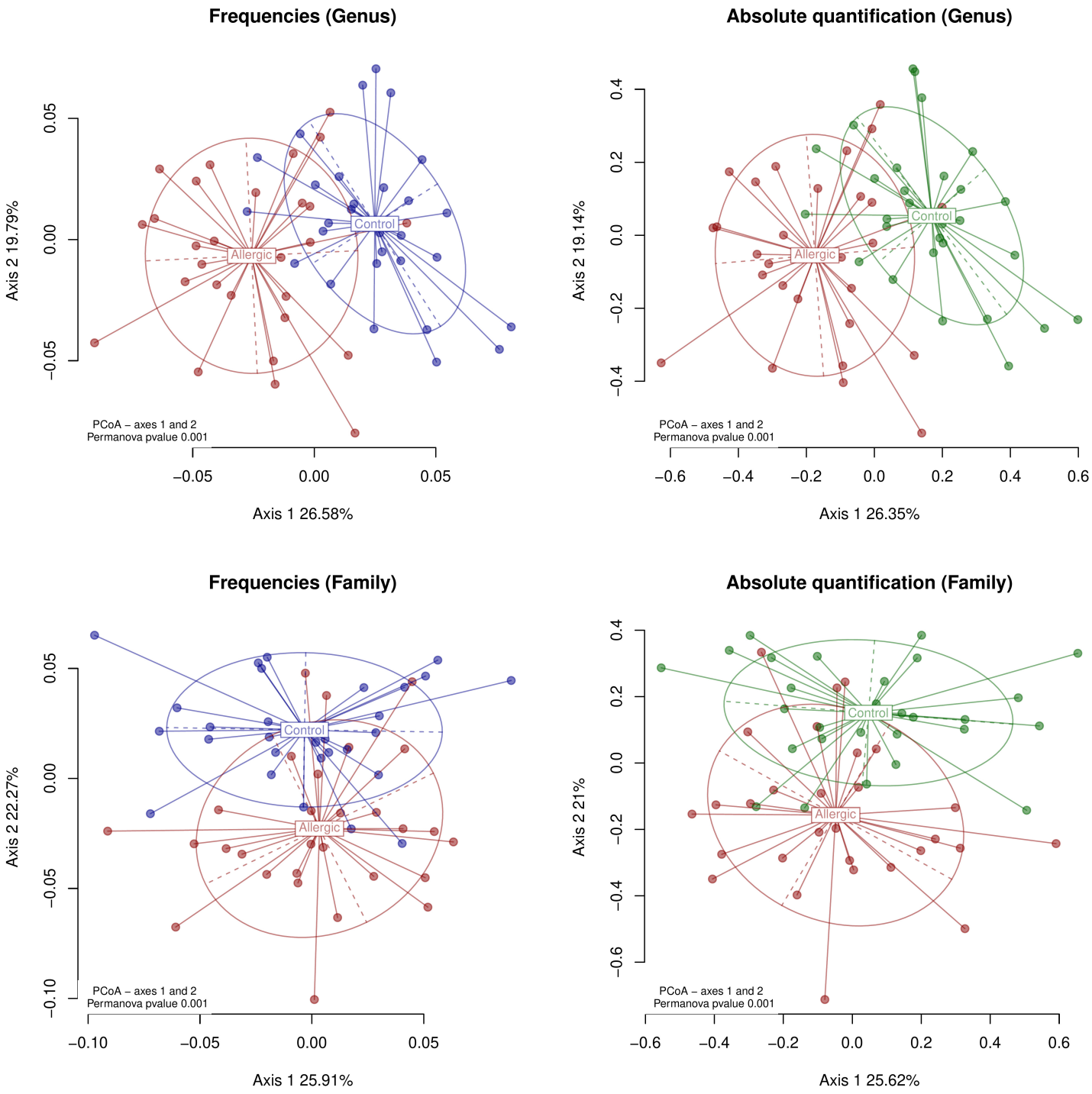


Figure 2

β-diversity analysis.

The microbiota composition of allergic vs control children were compared by Principal Coordinates Analysis (PCoA). (a-b) PCoA at the taxonomic level of the genus; (c-d) PCoA at the taxonomic level of the family. PCoA obtained for relative abundances are shown in plots a and c (samples from allergic children

are colored in dark blue and those from controls in light blue), while those computed from absolute abundances are shown in plots b and d (allergic in dark green and controls in light green). Allergic vs control microbiota compositions were compared using the PERMANOVA test and significant threshold was set < 0.05.

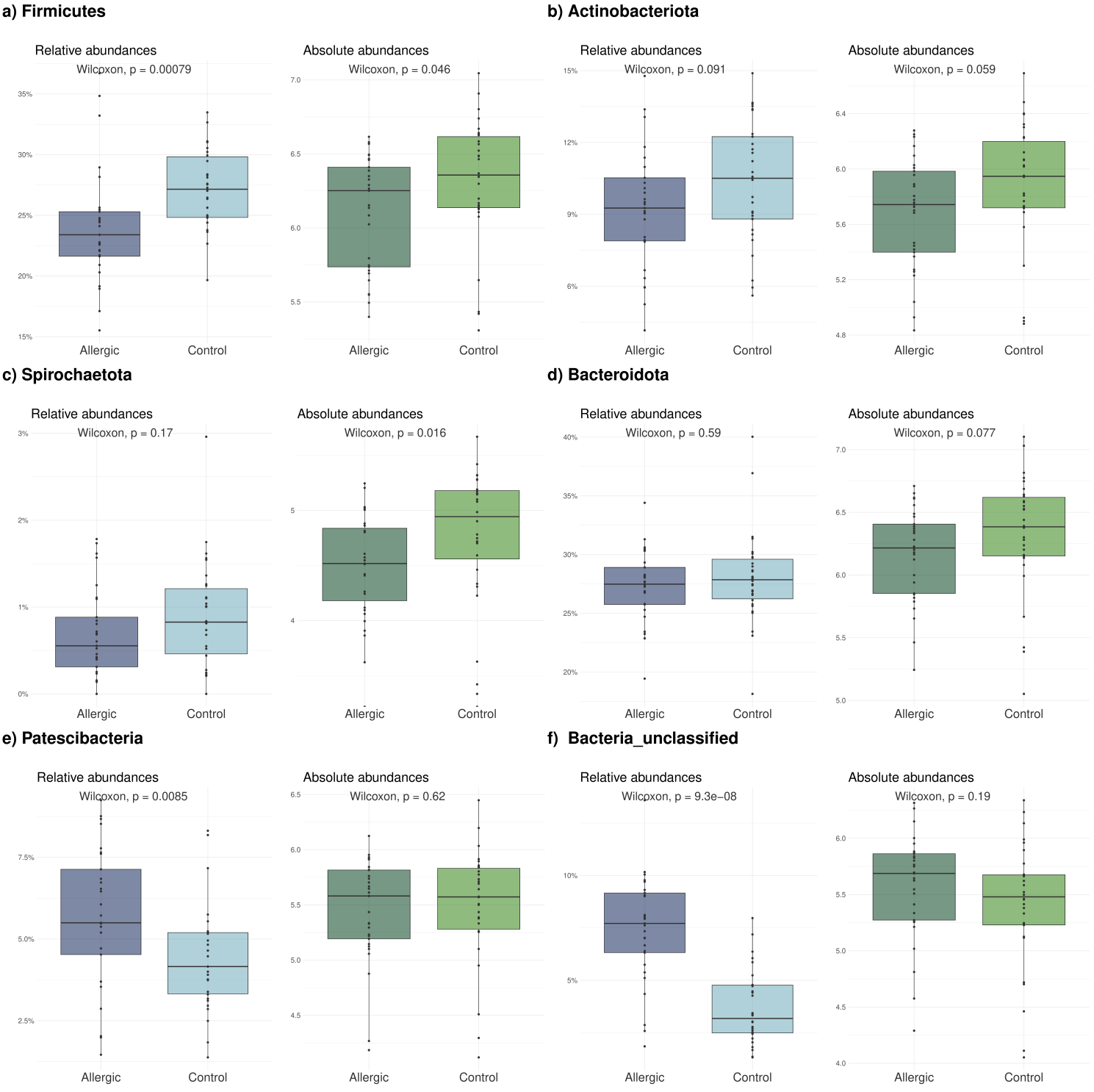


Figure 3

Absolute versus relative abundances of bacterial phyla

Plots a-f show relative and absolute abundances of bacterial phyla with significant differences in relative and/or absolute abundances between control and allergic children. In each plot, two couples of boxplots are shown: on the left, boxplots comparing the relative abundances between allergic (dark blue) and control (light blue) subjects; on the right, the boxplots comparing the absolute abundances between allergic (dark green) and control (light green) subjects.

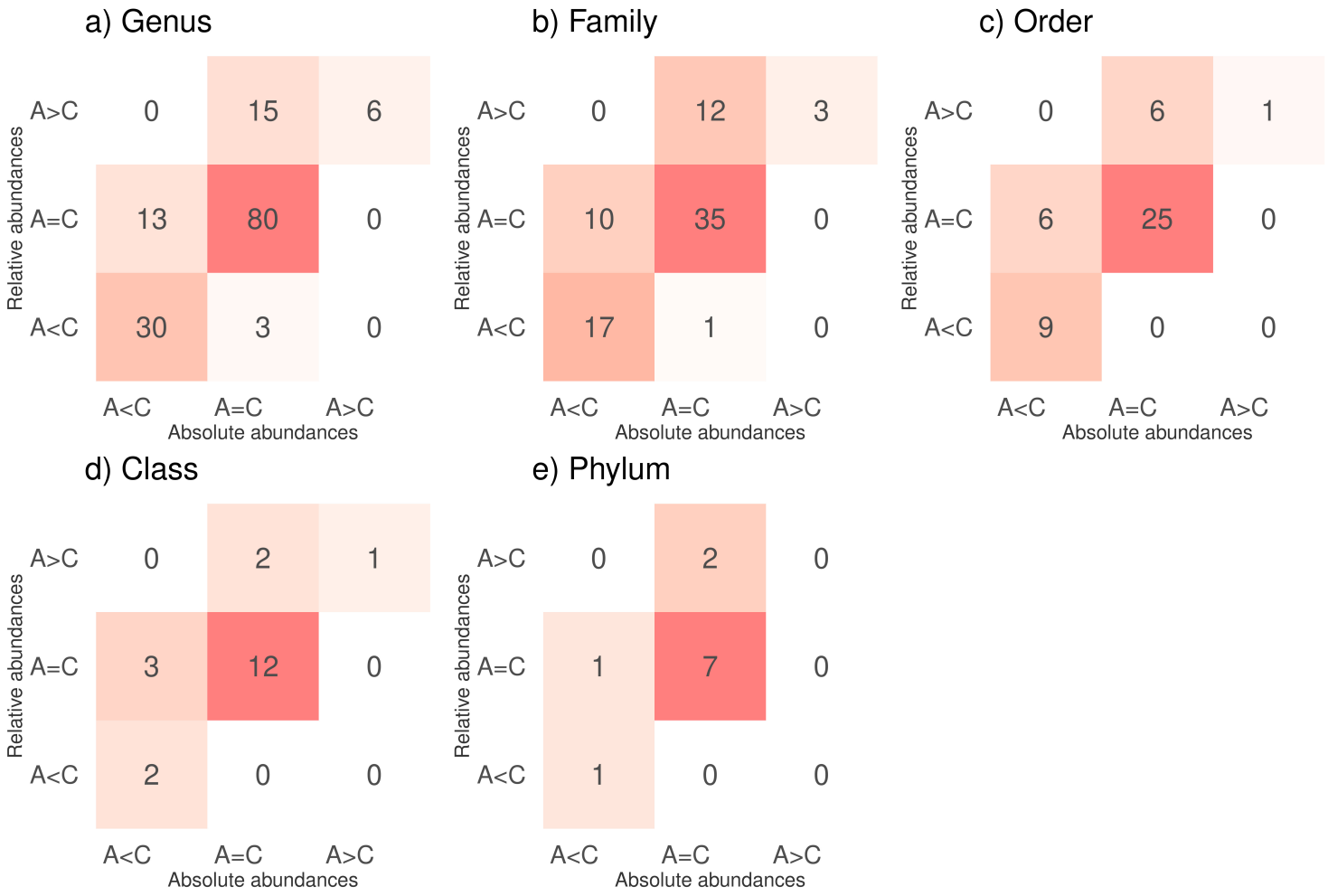


Figure 4

Concordance between absolute and relative abundances

Heatmap showing patterns of concordance and discordance between relative and absolute abundances across all taxonomic levels. Rows represent relative abundance comparisons, and columns represent absolute abundance comparisons. Row and column labels refer to comparisons between allergic (A) and control (C) children: A > C indicates taxa with statistically significantly higher abundance in allergic individuals; A = C indicates no statistically significant difference; A < C indicates taxa with significantly lower abundance in allergic individuals compared with control individuals. Values within cells indicate the number of taxa exhibiting the pattern defined by each row–column combination.

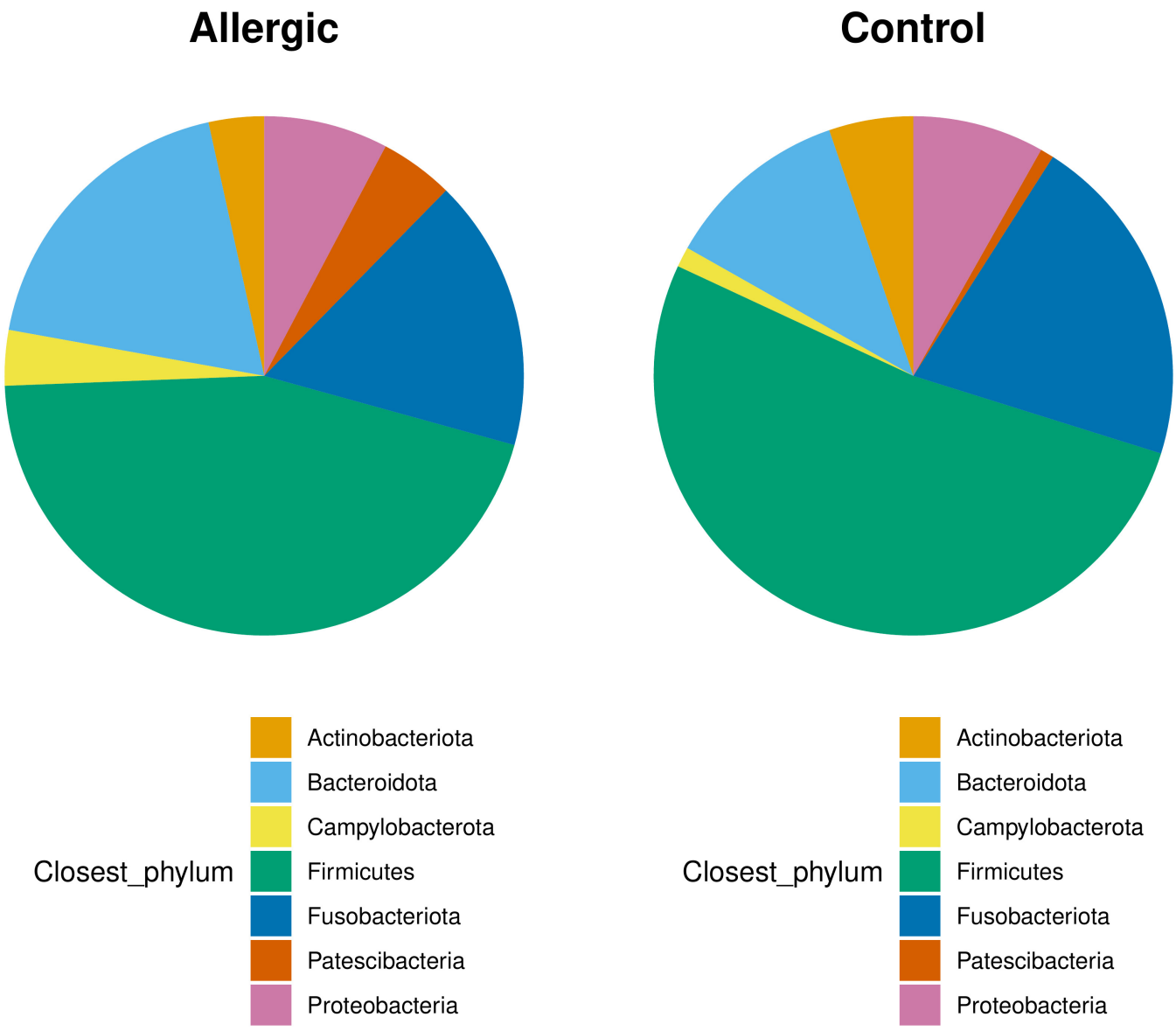


Figure 5

Taxonomic assignment of unclassified bacteria

Distribution of the bacterial phyla closest to sequences unclassified at the phylum level following 16S rRNA V3-V4 amplicon sequencing. The allergic group is shown on the left, and the control group on the right.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigures.pdf
- TableS1.xlsx