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.

Identification of the composition, stability, and origin of blood microbiome in humans

Maija Rozenberga
Latvian Biomedical Research and Study Centre
Rihards Saksis
Latvian Biomedical Research and Study Centre
Ilze Elbere
Latvian Biomedical Research and Study Centre
Liga Birzniece
Latvian Biomedical Research and Study Centre
Monta Briviba
Latvian Biomedical Research and Study Centre
Ilze Konrade
Riga Stradins University
Janis Klovins (klovins@biomed.lu.lv)
Latvian Biomedical Research and Study Centre

Research Article

Keywords: blood microbiome, circulating microbiome, 16S rRNA gene, microbiome origin, microbiome stability

Posted Date: September 2nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2019063/v1

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

Abstract

Background

An increasing number of research studies observe that human blood is not a completely sterile environment and has its own representative microbiome. This study aimed to determine the blood microbiome's composition, potential origin, and dynamics in humans.

Results

To detect the origin of exogenous bacterial nucleic acids in the blood, we determined taxonomic composition based on 16S rRNA gene analysis in samples obtained from skin, vaginal, oral, and faecal swabs along with whole blood samples in a group of 10 volunteers. We observed a presence of bacterial DNA with variable taxonomic composition in all blood samples strongly dominated by members of the *Pseudomonas* genus. In addition, we detected identical bacterial Amplicon Sequence Variants (ASV) that overlapped between blood and other locations in all participants. Overall, 27.4% median of all ASVs from blood were also found in various locations, with the highest number found in the samples collected from skin swabs. Overall, 25.3% of the ASV found in blood overlapped between the baseline and three-month blood samples, indicating the blood microbiome's relative stability.

Conclusions

We have presented for the first time a remarkable overlap between the bacterial composition of blood and other locations of the same individuals, allowing us to propose the skin microbiota as the primary source of blood-related exogenous DNA. Furthermore, our results add a piece of new knowledge on the stability of the blood microbiome, providing the basis for future studies to identify the potential effect of the blood microbiome on the phenotype or disease.

Background

Different microbial communities are found almost everywhere in and on the human body. However, commensal microbiome research is mainly focused on key sites such as the gut, skin, mouth, and vagina, and for a long time, human blood has been perceived as a sterile environment. This is because the presence of any microorganisms in blood was historically associated only with infectious diseases. However, new sequencing and culturing approaches have revealed some proof on the presence of a specific microbial composition also in the blood of healthy individuals [1,2].

Several studies have suggested that the blood microbiome potentially originates from other locations, particularly in cases of gastrointestinal tract diseases. Moreover, microbial signatures in the blood microbiome have been suggested as valuable biomarkers for diseases such as hypertension, type 2 diabetes, and chronic kidney disease, where increased intestinal permeability is observed [3–5]. Nevertheless, the likely source of a healthy human blood microbiome is still unclear. In principle, the blood

microbiome is established either by vertical or horizontal transmission (or a combination of both). Despite separate circulatory systems, vertical microbial transmission *in utero* might occur from mother to fetus, though the more convincing hypothesis suggests a potential to acquire the microbiota via horizontal transmission where translocation of bacteria from richer niches, most likely from minor skin or mucosa injuries during daily activities contributes to the formation of blood microbiome [1,2].

It is also still unclear whether the genetic material found in blood represents viable, non-viable, or dormant bacteria or only their remains, as current studies provide somewhat controversial results on the viability measures [1,6–8]. Blood is certainly a low microbial biomass microbiome sample, and it is very sensitive to contamination from the environment and laboratory reagents. Therefore, special attention should be paid to the inclusion and analysis of negative controls in all sample treatment steps to avoid false-positive results [1,9,10]. The number of studies focusing on the characterization of the blood microbiome is increasing. However, many aspects of this microbiome subpopulation, including the stability of the blood microbiome composition over time remain elusive. In this study, we aimed for the first time to analyse the composition, potential origin, and longitudinal dynamics of the blood microbiome in a pilot group of volunteers from the general population.

Methods

Study design and sample collection

A total of 10 individuals from the general population were enrolled in the study, meeting the following inclusion criteria – age 18–64 years, no antibiotics used in the last 2–3 months as well as no diarrhoea seven days before involvement in the study. Samples and information on all participants were included in the Genome Database of the Latvian Population (LGDB), and recruitment was organised in accordance with previously established practices [11].

Participants were invited to two visits: whole blood samples, swabs from the skin (scalp, upper back, volar forearm), vagina, oral mucosa, and faeces were collected during the first visit; while whole blood samples and faecal swabs were collected again three months later (Fig. 1). Along with the sample collection, data about the date and time of sample collection, sex, age, height, weight, dietary lifestyle, mode of delivery, smoking and medical history, and medication consumption during the previous two months were collected. Whole blood samples from the median cubital vein were collected by medical personnel after disinfection of the sample collection site; participants were in a fasting state. To avoid microbial DNA contamination from the skin during the blood collection procedure, the blood samples were drawn in three EDTA-containing blood tubes, where either the 2nd or 3rd tube was selected for the subsequent 16S rRNA gene analysis. The samples were stored at +4 ºC degrees until further processing. Microbial DNA from blood was isolated within 1–3 days of sample collection.

Oral and skin swabs were collected by medical personnel; vaginal and faecal swabs were collected by participants at home. No food was allowed for 30 minutes before oral microbiome sampling. Vaginal

samples from women of reproductive age were not collected during menstruation. eNAT® System with nucleic acid preserving solution (Copan, USA) was used for swab-based collection. Two aliquots were obtained from each sample type. After collection, the samples were stored at room temperature and transported to the laboratory as soon as possible, but no later than within 24 hours, where they were stored at -80°C until the microbial DNA was isolated.

Sample preparation

Microbial DNA was isolated from the whole blood using the phenol-chloroform method in accordance with the standard operation protocol developed by the LGDB [11]. To reduce contamination, in addition to standard laboratory procedures, sample processing was performed using a surgical face mask and in a laminar flow cabinet.

When all samples from one individual were obtained, microbial DNA from swabs was extracted under sterile conditions in a Biosafety Level 2 cabinet. 200 µl of eNAT® buffer (Copan, USA) containing collected microbial nucleic acids was used to isolate the microbial DNA using FastDNA Spin Kit for Soil (MP Biomedicals, Santa Ana, CA, USA) and FastPrep Instrument following the instructions of the manufacturer. DNA concentrations were evaluated using a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). Extracted microbial DNA was validated by agarose gel electrophoresis.

NGS sequencing

The V3-V4 hypervariable region of the bacterial 16S rRNA gene was PCR-amplified using Phusion U Multiplex PCR Master Mix (Thermo Scientific, USA) and a 341F/805R primer pair. Embedded sample identification sequences were added during a second PCR using appropriate oligonucleotides (see Table S1 for the list of oligonucleotides (Additional File 1)). At least one blank control was prepared for every 4–6 samples in the first amplification, and at least one empty control was prepared for every 6–8 samples in the second amplification. Successful amplification of samples and purity of negative controls were determined by agarose gel electrophoresis. The PCR amplicons were purified using NucleoMag magnetic beads (Macherey-Nagel, Düren, Germany), and their quantity and quality were evaluated with the Agilent 2100 Bioanalyzer DNA High Sensitivity chip (Agilent Technologies, Santa Clara, CA, USA).

Sample sequencing was performed on the Illumina MiSeq platform using the MiSeq Reagent Kit v2 (500-cycles) (Illumina, USA), obtaining at least 100 thousand sequencing reads per sample. All samples from each study participant were placed on the same chip to prevent a batch effect.

Extensive use of negative controls (NC) was ensured in all steps (including blank controls when starting sample preparation with a new set of reagents) and working areas. In addition, to ensure sterility, working areas, instruments, and individual reagents were disinfected with UV and disinfectants at all stages, and the work was performed in protective clothing.

Data analysis

Raw sequencing data were first processed by performing a quality check with FastQC(*v0.11.9*) [12] and MultiQC(*v1.10*) [13]. Subsequently, data were read into QIIME2(*v2021.2*) software, where most of the remaining analysis was done.

To extract V3-V4 hypervariable region, the data were trimmed with the QIIME2 [14] Cutadapt plugin [15] by using the 341F and 805R primer sequences. To check for sequence length and quality after trimming, the Demux plugin was used. Next, data were denoised using the DADA2 plugin [16], where the truncation length parameters for forward and reverse reads were set as the second percentile of read length distribution while accounting for the recommended 12 base pair overlap to maximise the number of successfully merged sequences. Data quality trimming was not performed before hypervariable region extraction and denoising to take advantage of DADA2 capabilities of learning from sequencing error rates to make more accurate inference of Amplicon Sequence Variants (ASV). As a result, an ASV table was obtained and further filtered by frequency of five sequences per ASV in at least one sample, to remove possible technical errors of library preparation and sequencing.

To remove possible bacterial contamination of samples during library preparation and sample handling, a frequency filtered ASV table was exported to R(*v4.1.0*) where Decontam(*v1.12.0*) package was used on phyloseq(*v1.38.0*) format data to detect the contaminants [17,18]. First, the package was run in prevalence mode with the default threshold (0.1) to obtain the frequency values of all the ASVs (N = 4342) both in true samples and blank controls. Mode values were calculated for all true sample and blank control subset groups [19], and ASVs with frequency values lower than the calculated mode were selected. From the remaining ASVs, mode of frequency values was calculated once again and used as the subset group-specific threshold value for the Decontam prevalence test. A list of contaminant ASVs was obtained from all the subgroups and excluded from the filtered ASV table using the QIIME2 feature-table plugin. No manual filtering of contaminant features was performed, as it would introduce a bias in regard to the chosen source of contaminant reference as it can not be 100% comprehensive.

To inspect how sequencing depth affects the Operational Taxonomic Unit (OTU), sample count, and Shannon diversity index values, alpha rarefaction curves were calculated and visualised using the QIIME2 diversity plugin by providing the maximum value of observed features (84545) in a sample from the feature table.

To calculate diversity indexes for the decontaminated data, a feature subset with a size 965 of all ASVs from all true samples was made by visually inspecting ASV count distribution in the feature table and determining an arbitrary cut-off threshold with a goal to preserve as many samples as possible while selecting for the largest feature subset value possible. Accordingly, to calculate and visualise both phylogenetic and non-phylogenetic diversity metrics on the feature subsample, the diversity plugin was used.

Sequences were then classified using the naïve Bayes classifier by following the RESCRIPt [20] workflow while using the SILVA(*v138.1*) small subunit non-redundant database with a 99% identity as a reference [21]. Data from the database were pre-processed to optimise for the highest possible classification confidence. First, only the Archae, Bacteria, and Eukaryote domains were selected with respective minimal sequence lengths for each domain at 900, 1200, and 1400. Next, the contents of the database were dereplicated to preserve only unique taxa. The V3-V4 region was extracted by providing the 341F and 805R primer sequences. Dereplication was carried out again, and finally, a classifier was trained on the cleaned database.

A bar plot containing each sample and sample group was made to visually assess and compare the taxonomic diversity of said sample groups. Lastly, the phyloseq package was again used to construct additional bar plots for specific sample groups and to construct an NMDS plot with a confidence interval of 95%.

To assess the effects of decontamination, all the steps mentioned above, except for the decontamination and the NMDS plot, were performed on the non-decontaminated data as well. It was subsampled to the same 1339 ASVs to match the sequencing depth of the decontaminated data. Values for the Shannon, Faith's phylogenetic diversity, Pielou's evenness, UniFrac Weighted, and Unweighted Beta diversity indexes were exported and for each body location compared to those of the decontaminated data using either the Student's t-test or the Wilcoxon test depending on whether values for certain body locations differed significantly from the normal distribution.

To assess the stability of the samples over time, the Galaxy website for statistical analysis of biomedical research data was used (https://usegalaxy.org/). It uses the LefSE analysis tool, which explains differences between groups using the linear discriminant analysis (LDA) effect size (LEfSe) algorithm [22]. For all tests, the significance threshold = 0.05 and the LDA threshold = 2.0.

Results

Characteristics of the study group and samples

Seven female and three male volunteers without acute illness participated in this study. Median age of the study group was 40 (interquartile range (IQR) = 21) years, while body mass index (BMI) was 26 (IQR = 4.3) kg/m^2. All details about each participant can be found in Table S2 (Additional File 1).

The median number of reads obtained per sample was 83534.50 (IQR = 29850.75). The depth of sequencing and the used database allowed to classify microorganisms mainly up to the genus level and up to species level in some cases. A total of 4391 ASVs were detected in all samples, including 4287 ASVs in true samples and 210 ASVs in controls – 28 phyla and 460 genera. After decontamination 4240 ASVs remained in total (28 phyla and 454 genera). To provide reliable results, 11 negative controls were included. These control samples were dominated by bacterial genera such as *Ralstonia* (39.1%),

Hydrogenophilaceae_uncultured (25.4%), *Undibacterium* (13.7%), *Curvibacter* (5.8%), *Thermus* (4.4%), *Acidovorax* (4.2%), *Lactobacillus* (2.6%), *Enhydrobacter* (1.3%). Contaminant ASVs are listed in Table S3 (Additional File 1).

Alpha and beta diversity

To characterise the microbiome from various locations, we first evaluated alpha diversity using the Shannon entropy index and beta diversity using PCoA (Fig. 2). Comparing the alpha diversity indexes, no significant difference was observed between blood samples from both visits, but it was observed when comparing V1 blood samples to other locations (Fig. 2A). As expected, oral and faecal samples displayed distinctly higher alpha diversity index than other locations. The median Shannon index of blood samples was 2.29 (IQR = 1). A graphical representation of the diversity of Pielou's evenness and Faith's PD indexes is shown in Fig. S1, S2.

PCoA analysis showed the grouping of the samples according to the collection site, with the most distinct grouping of faecal samples. Based on this analysis, blood samples overlapped with the samples from different skin locations. Beta diversity of blood samples alone depicted substantial similarity between the two blood samples from each individual (Fig. 2B, 2C, 2D). Boxplots representing weighted and unweighted UniFrac distance value distributions for each location are shown in Fig. S3, S4.

Blood microbiome composition

An overall characterisation of the taxonomic profile in all collected blood samples is depicted in Fig. 3. At the phylum level, *Proteobacteria* predominated in almost all V1 and V2 blood samples (91.8%), followed by *Firmicutes* (4.5%), and *Bacteroidetes* (1.7%) (Fig. 3A). *Pseudomonas* was the most frequent genus (86.1%), while *Veillonella* (3.2%), *Prevotella_7* (1.6%), and others were less common (<1.5 %) (Fig. 3B).

Microbiome composition at other locations at the genus level is shown in Fig. S5 – S11 (Additional File 2).

Origin and stability of blood microbiome

To assess the possible origin of bacterial DNA found in the blood microbiome, we searched for the presence of exact ASVs found in the V1 blood samples and in other samples collected at the same time-point from the same individual (Table 1). Depending on the participant 10.3–35.0% of ASVs found in blood were also identified in the samples from other locations. We also observed 5.5–36.4% ASV overlap between pairs of participant-specific blood samples taken at a three-month interval (Table 1). The highest number of ASVs identical between V1 blood and other locations was found in the samples collected from skin swabs. In addition, from 0.39% to 0.83% of all ASVs found in samples collected from different

locations were also identified in blood samples (Table 2). Many ASVs found to be identical in blood and other locations were identified in more than one participant. However, a number of the observed ASVs were unique at the individual level, as we observed 32 ASVs overlapping between blood and a single location in a single individual (Table 2). No ASVs were found to overlap between blood and several locations in a single participant.

Table 1. Observed overlap of ASVs between V1 blood samples, other V1 samples, and V2 blood samples.

Participant	Number of ASVs in V1 blood samples	Number of ASVs both in V1 blood sample and at least one other location of the same individual (%)	The proportion of ASVs from V1 blood samples as a percentage of all ASVs observed in the individual (%)	Number of ASVs in V2 blood samples	Overlap between V1 and V2 blood samples (%)
S1	18	5 (27.78)	1.77	28	9 (24.32)
S2	20	5 (25.00)	2.36	25	12 (36.36)
S3	29	3 (10.34)	3.40	21	11 (28.21)
S4	19	6 (31.58)	2.31	23	3 (7.69)
S5	26	7 (26.92)	3.49	22	10 (26.32)
S6	18	6 (33.33)	1.81	22	11 (37.93)
S7	44	7 (15.91)	4.63	45	11 (14.10)
S8	20	7 (35.00)	3.14	18	2 (5.56)
S9	20	6 (30.00)	4.04	8	2 (7.69)
S10	40	9 (22.50)	6.49	23	14 (28.57)
Median	***20***	***6 (27.35)***	***3.27***	***22.5***	***10.5 (25.32)***
IQR	***9***	***1.75 (8.06)***	***1.58***	***3.25***	***6.5 (19.19)***

IQR: interquartile range, ASV: amplicon sequence variant. V1: first visit, V2: second visit after three months.

Table 2. Observed overlap between V1 blood samples and other V1 samples representing specific microbiome subpopulations.

Location	Number of ASVs concurrent with V1 blood samples (%)	OTUs representing unique ASVs identical between blood and the particular location found in a single participant (number of different ASVs)	OTU representing ASVs identical between blood and the particular location found in more than one participant (number of different ASVs)
Oral mucosa	23 (0.53)	*Lawsonella (1); Rothia (1); Veillonella (1)*	*Lawsonella (1); Pseudomonas (3); Ralstonia (1); Rothia (1); Streptococcus (2); Veillonella (1)*
Scalp	31 (0.71)	*Acinetobacter (1); Brevundimonas (1); Micrococcus (1); Paracoccus (1); Streptococcus (1); Undibacterium (1)*	*Acinetobacter (1); Bifidobacterium (1); Brevundimonas (1); Lawsonella (2); Micrococcus (1); Paracoccus (1); Pseudomonas (3); Ralstonia (2); Staphylococcus (2); Streptococcus (1); Undibacterium (1)*
Upper back	36 (0.83)	*Acidovorax (1); Bifidobacterium (1); Lawsonella (1); Micrococcus (1); Paracoccus (1); Rothia (1); Sphingomonadaceae_Unclassified (1); Streptococcus (1); Undibacterium (1); Veillonella (1)*	*Undetermined (1); Acidovorax (1); Bifidobacterium (1); Brevundimonas (1); Lawsonella (1); Micrococcus (1); Paracoccus (1); Pseudomonas (4); Ralstonia (2); Rothia (1); Staphylococcus (3); Streptococcus (1); Undibacterium (1); Veillonella (1)*
Volar forearm	38 (0.88)	*Bifidobacterium (1); Chloroplast (1); Corynebacterium (1); Paracoccus (1); Rothia (1); Streptococcus (1); Sutterella (1); Veillonella (1)*	*Bifidobacterium (1); Brevundimonas (1); Chloroplast (1); Corynebacterium (1); Paracoccus (1); Pseudomonas (2); Ralstonia (2); Rothia (1); Staphylococcus (3); Streptococcus (1); Sutterella (1); Undibacterium (1); Veillonella (1)*
Vagina	19 (0.44)	*Prevotella_7 (1); Staphylococcus (1)*	*Pseudomonas (4); Ralstonia (1); Staphylococcus (1); Undibacterium (1)*
V1 faeces	22 (0.51)	*Bacteroides (1); Blautia (1); Sutterella (1)*	*Bacteroides (1); Bifidobacterium (2); Blautia (1); Pseudomonas (2); Ralstonia (1); Sutterella (1)*

IQR: interquartile range, ASV: amplicon sequence variant, V1: first visit, OTU: operational taxonomic unit.

To evaluate the stability of the blood microbiome, we used LEfSe algorithm to compare samples collected during the first visit with samples collected after three months. No ASVs and OTUs with significantly different abundance between the first and second visit's blood samples were found and the first and second visit's faecal samples. Observed ASVs partially overlapped between the longitudinally collected whole blood samples (Table 1).

Discussion

There is a growing body of evidence in the scientific literature suggesting the presence of bacteria in the human blood, not only in septic patients but also in healthy individuals. There are various hypotheses about the potential origin of the blood microbiome, but it has not been evaluated in detail. This study determined the presence of bacterial DNA in various samples from the same individual by analysing the 16S rRNA gene amplicons.

In samples with low and very low microbial biomass, contamination can account for up to 90% of the resulting reads; therefore, the inclusion of negative controls at different stages of sample processing and their analysis is very critical [23]. Due to the relatively recent focus on the importance of negative controls in studies with low microbial biomass samples, many studies lack appropriate negative and blank controls subjected to sequencing. This problem has generally affected microbial research, giving a false view of the microbiome composition of different low biomass samples [24]. For instance, previous studies have shown that the placenta has its unique microbiome [25,26]. Olomu et al., on the contrary, have shown that after correcting for *kitome* and cross-contamination events, it was not possible to identify placenta-specific microbiome [9]. Therefore, we paid particular attention to the analysis of negative controls, including a number of negative controls from different working areas and reagents.

Previous studies that analyse the content of blank controls have identified both human microbiome-associated taxa, such as *Prevotella* and *Lactobacillus*, and members of the genus *Janthinobacterium* and *Thermus*, commonly found in water and hot springs typically not present in human-associated microbiomes [27,28]. Most commonly, bacteria of the genus *Ralstonia* are identified in the blank controls. These bacteria are found in plants, soil, and water and are typical laboratory contaminants [6,24,29,30]. On the other hand, this genus cannot be entirely excluded from data since *Ralstonia* is often associated with human opportunistic pathogenic microbiota, especially in hospital settings [29,31]. Furthermore, previous reports have found that members of the genus *Ralstonia* were associated with intestinal diseases [32] and cystic fibrosis [33]. In other studies, various reagent kits were tested and analysed, resulting in the identification of the *Pseudomonas* genus accounting for most of the contamination (>70%), followed by the *Burkholderia* genus (17%) [34], while Whittle et al. observed *Serratia* as a contaminant in the blank controls of blood microbiome study [6] and Gosiewski et. al. – *Sphingomonas* [35]. However, in many samples, the majority of taxonomic composition was formed by bacteria widespread in the environment, and it can therefore be part of the human microbiome in the locations that frequently interact with this environment.

The presence of nucleic acids from Gram-negative bacteria was observed in the blood, a strong dominance of *Proteobacteria* phylum, *Pseudomonas* genus, followed by *Firmicutes* and *Bacteroidetes*, similar to what has been found in other studies [3,6,36]. On the contrary, T. Gosiewski et al. observed a significant predominance of anaerobic bacteria, including Gram-positive *Bifidobacteriales* order, and decreased abundance of *Proteobacteria* phyla in healthy volunteers compared to patients with sepsis [35]. There are controversial observations in the literature about the *Pseudomonas* genus as it is

ubiquitous in the environment, including humans and animals, and plants [37]. It cannot be excluded that their presence in the blood samples may result from potential contamination [6,38] as the *Pseudomonas* were present in negative controls. On the other hand, due to their overwhelming presence in the environment, bacteria of this genus may predominate in very low biomass samples naturally, as this trend has been observed in other studies of both the lung microbiome [39] and the deeper layers of the dermis, which were previously considered sterile [40]. Several studies have also used the classical culture-based method to determine the presence and composition of a blood microbiota. Panaiotov et al. found that in healthy people, the microorganisms in the blood are in an inactive state, while under appropriate conditions, they may change their activity status, also confirming that the predominant phylum of bacteria in the blood is *Proteobacteria* [41]. Whittle and colleagues identified *Staphylococcus* genus by culturing blood samples, which more likely was the result of contamination from skin [6].

A strong dominance of the *Pseudomonas* genus was observed in most blood samples, but other microorganisms were also detected. Although in much smaller numbers, we also found *Streptococcus, Staphylococcus, Lactobacillus, Veillonella,* and other bacterial genera in the blood samples. These are typical members of oral, gut, vaginal, and skin microbiome and were found in different proportions in all these locations, while *Brevundimonas, Sphingomonadaceae_Unclassified* were only found in all skin locations. It is considered that microorganisms enter the circulation from different body parts, such as the gut [36]. This has been confirmed in the case of various gastrointestinal and other diseases in which intestinal permeability has been altered [4,42,43]. At the same time, Shah et al. later found that the blood microbiome does not directly reflect the gut microbiome [43]. Our study confirms this observation as the beta diversity analysis shows that there is a much larger distance between blood and faecal samples than between blood and skin samples, suggesting that the blood microbiome of healthy individuals may originate from the skin microbiome, which partly coincides with Whittle et al. results [6]. In the same study, it was observed that the blood microbiome is also similar to the oral microbiome [6]. In our study, bacterial taxa from blood mostly overlap with those from various other skin samples. Still, a significant fraction of bacteria found in the blood did not overlap with other locations. This may suggest that the blood has its own unique microbiome maintained independently for a longer time. However, it should be noted that due to practical reasons, some locations were not tested (e.g. respiratory tract), and also, the skin microbiome may contain a significantly larger number of different bacteria, as only a tiny fraction of the whole skin is tested even if several different locations were chosen.

Using ASVs to compare the presence of the same bacteria in different samples provides a much better resolution to track down the exact source of bacterial presence. We, therefore, searched for the presence of 100% identical ASVs between blood and other locations. It is not surprising that the large number of sequence variants identified as identical in a blood sample and other locations were present in several individuals, as they represented commonly found bacterial taxa. On the other hand, one may not exclude contamination in such cases even after applying stringent measures to filter out the possible contaminants. The finding of 32 unique ASVs that were identical between the blood and another location but each only in a single individual is strong proof that we indeed are able to prove the origin for at least some fraction of blood microbial sequences. It is important to note that we did not observe any unique

ASVs that would overlap between the locations of two different individuals acting as strong proof against possible cross-contamination events.

Paise et al. examined different fractions of blood, considering previous studies that suggested some bacteria may be dormant in red blood cells. They found significant differences between the taxonomic profiles of the different blood fractions, where potential blood pathogens that we also found, such as *Acinetobacter, Corynebacterium, Pseudomonas, Staphylococcus*, were relatively more abundant in the red blood cell fraction [36]. Therefore, whole blood samples were used in this study. Moreover, a feasibility study concluded that whole blood samples could provide a broader picture than individual blood fractions while also reducing the risk of microbial contamination due to the smaller number of laboratory interventions (*unpublished data*). In addition, since blood is considered to have a particularly low level of microbial biomass and thus is more sensitive to contamination, a different DNA extraction protocol was used for blood compared to the other samples. To extract the maximum amount of microbial DNA, whole blood from EDTA tubes was treated with the phenol/chloroform extraction method.

Interestingly, non-human-associated microbiome taxa were also observed in the blood in this study. *Chloroplast* genus was found on the forearm and in the blood sample of one individual. Most likely, this bacterial genus, usually associated with plants [44], entered the bloodstream from a forearm injury. Another study has found the presence of a non-human microbiome genus in individuals regularly using herbal cosmetics [34]. In some studies, OTUs and bacteria classified as mitochondria, chloroplasts, or eukaryotes are filtered out from the outset [45]. Almost all chloroplast DNA identified in our study was present only in skin samples.

Alpha diversity of taxa in blood samples is relatively low, with a median Shannon index of 2.14, a value that is similar to the findings of other studies [3]. An equal index value was also calculated for skin samples, while it was 3 times higher for faecal samples.

To determine the stability of the microbiome over time, blood and faecal samples were collected twice with a 3-month interval. The linear discriminant analysis (LDA) effect size (LEfSe) algorithm did not show significant differences in either the ASV or OTU data, indicating that the blood and gut microbiome of these individuals remained stable during the three months. The major bacterial taxa of the blood microbiome overlapped between the samples from both visits, which indicates that the overall blood microbiome is stable in healthy subjects. Microbiome stability is also confirmed by other studies as the composition of the intestinal microbiome, as well as skin microbiome in healthy people, can persist for years [46,47], with certain fluctuations under the influence of certain factors over time [48].

Some limitations of our study should be mentioned. First, we had a relatively small sample size and rather a heterogeneous study group with respect to age, sex, and BMI. However, our primary goal was not to associate specific blood microbiome composition with phenotype but evaluate the origin of the blood microbiome compared with the microbiome from other locations. For this purpose, having a phenotypically heterogeneous group may be seen as an advantage. Secondly, different methods were used to isolate microbial DNA from blood samples and swabs, and the DNA extracted from blood

samples contained a large amount of human DNA. Therefore, we filtered out those ASVs that were present in greater numbers in control samples than in biological samples, according to Karstens et al. paper [23], while others have excluded overlapping sequences between the samples and controls [41]. We chose this method so that the taxa that are characteristic of the human microbiome would not be filtered out, but as a result, the data used in the analysis include some taxa that are usually not associated with the human microbiome. We also did not estimate the viability of bacterial cells. It is therefore impossible to determine whether the genomic sequences detected in the blood come from live bacteria or whether only their genetic material is present in the circulation.

Conclusions

To the best of our knowledge, this is the first study presenting results on blood microbiome origin by analysing samples collected from multiple locations from the same individuals at the same time point. We found that the blood microbiome of healthy individuals is predominated by members of the genus *Pseudomonas* of the *Proteobacteria* phylum and has the highest similarity with the skin microbiome. Furthermore, the composition of the blood microbiome remains stable for at least three months.

Despite the small sample size, this study depicts results that can be further used in designing and interpreting other studies. In addition, the results and findings of this study will complement the existing knowledge on the blood microbiome and will serve as a basis for future research on the blood microbiome in the context of various diseases.

Abbreviations

16S rRNA gene: 16S ribosomal RNA gene

ASV: Amplicon sequence variant

BMI: Body mass index

IQR: Interquartile range

OTU: Operational taxonomic unit

PCR: Polymerase chain reaction

V1: First visit

V2: Second visit after three months

Declarations

Ethics approval and consent to participate

This study has received the permission of the Central Medical Ethics Committee of Latvia (No. 01-29.1 / 6359). All participants provided signed consent.

Consent for publication

Not applicable.

Availability of data and materials

The dataset supporting the conclusions of this article is available in the Sequence Read Archive repository under the accession project number PRJNA874911 (https://www.ncbi.nlm.nih.gov/bioproject/874911).

Competing interests

The authors declare that they have no competing interests.

Funding

The research was carried out within the framework of the project “Investigation of the origin and dynamics of human blood microbiome, and its role in chronic disease”. Project identification No.: lzp-2019 / 1-0116.

Maija Rozenberga was supported by University of Latvia Foundation with the Doctoral scholarship "MikroTik" in the field of exact and medical sciences.

Contributions

MR carried out participant selection, sample preparation and protocol optimisation, data interpretation and wrote the original manuscript draft. RS performed data analysis, interpretation and validation, and manuscript writing and editing. IE participated in the study design, data interpretation, manuscript writing and editing, and supervision. LB carried out sample processing and library sequencing. MB participated in the study design and manuscript editing. IK participated in the study design and participant selection. JK participated in the study design, funding acquisition, data interpretation, manuscript editing, and supervision. All authors read and approved the final manuscript.

Acknowledgements

The authors thank all volunteers for their participation in the study and as well as Genome Database of Latvian Population and Latvian Biomedical Research and Study Centre for the organization of participants recruitment and biological sample processing and data storage.

References

1. Castillo DJ, Rifkin RF, Cowan DA, Potgieter M. The healthy human blood microbiome: Fact or fiction? Front Cell Infect Microbiol. Frontiers Media S.A.; 2019;9.
2. Velmurugan G, Dinakaran V, Rajendhran J, Swaminathan K. Blood Microbiota and Circulating Microbial Metabolites in Diabetes and Cardiovascular Disease. Trends in Endocrinology and Metabolism. Elsevier Inc.; 2020. p. 835–47.
3. Qiu J, Zhou H, Jing Y, Dong C. Association between blood microbiome and type 2 diabetes mellitus: A nested case-control study. J Clin Lab Anal. John Wiley and Sons Inc.; 2019;33:e22842.
4. Shah NB, Allegretti AS, Nigwekar SU, Kalim S, Zhao S, Lelouvier B, et al. Blood microbiome profile in CKD: A pilot study. Clinical Journal of the American Society of Nephrology. American Society of Nephrology; 2019;14:692–701.
5. Jing Y, Zhou H, Lu H, Chen X, Zhou L, Zhang J, et al. Associations Between Peripheral Blood Microbiome and the Risk of Hypertension. Am J Hypertens. Oxford Academic; 2021;34:1064–70.
6. Whittle E, Leonard MO, Harrison R, Gant TW, Tonge DP. Multi-Method Characterization of the Human Circulating Microbiome. Front Microbiol. Frontiers Media S.A.; 2019;9:3266.
7. Schierwagen R, Alvarez-Silva C, Madsen MSA, Kolbe CC, Meyer C, Thomas D, et al. Circulating microbiome in blood of different circulatory compartments. Gut. 2019;68:578–80.
8. Potgieter M, Bester J, Kell DB, Pretorius E. The dormant blood microbiome in chronic, inflammatory diseases. FEMS Microbiol Rev. Oxford University Press; 2015. p. 567–91.
9. Olomu IN, Pena-Cortes LC, Long RA, Long RA, Vyas A, Krichevskiy O, et al. Elimination of “kitome” and “splashome” contamination results in lack of detection of a unique placental microbiome. BMC Microbiol. BioMed Central; 2020;20.
10. Schierwagen R, Alvarez-Silva C, Servant F, Trebicka J, Lelouvier B, Arumugam M. Trust is good, control is better: Technical considerations in blood microbiome analysis. Gut. BMJ Publishing Group; 2020. p. 1362–3.
11. Rovite V, Wolff-Sagi Y, Zaharenko L, Nikitina-Zake L, Grens E, Klovins J. Genome Database of the Latvian Population (LGDB): Design, Goals, and Primary Results. J Epidemiol. J Epidemiol; 2018;28:353–60.
12. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Internet]. 2010. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
13. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. Oxford Academic; 2016;32:3047–8.

14. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 2019 37:8. Nature Publishing Group; 2019;37:852–7.
15. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. EMBnet Stichting; 2011;17:10–2.
16. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 2016 13:7. Nature Publishing Group; 2016;13:581–3.
17. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One. Public Library of Science; 2013;8:e61217.
18. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al. Orchestrating high-throughput genomic analysis with Bioconductor. Nature Methods 2015 12:2. Nature Publishing Group; 2015;12:115–21.
19. Davis NM, Proctor DiM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome. BioMed Central Ltd.; 2018;6:1–14.
20. Robeson MS, O'Rourke DR, Kaehler BD, Ziemski M, Dillon MR, Foster JT, et al. RESCRIPt: Reproducible sequence taxonomy reference database management for the masses. bioRxiv. Cold Spring Harbor Laboratory; 2020;2020.10.05.326504.
21. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Res. Oxford Academic; 2014;42:D643–8.
22. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol. Genome Biol; 2011;12.
23. Karstens L, Asquith M, Davin S, Fair D, Gregory WT, Wolfe AJ, et al. Controlling for Contaminants in Low-Biomass 16S rRNA Gene Sequencing Experiments. mSystems. American Society for Microbiology; 2019;4:290–309.
24. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends Microbiol. Elsevier Ltd; 2019. p. 105–17.
25. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. The placenta harbors a unique microbiome. Sci Transl Med. American Association for the Advancement of Science; 2014;6:237ra65-237ra65.
26. Bassols J, Serino M, Carreras-Badosa G, Burcelin R, Blasco-Baque V, Lopez-Bermejo A, et al. Gestational diabetes is associated with changes in placental microbiota and microbiome. Pediatr Res. Nature Publishing Group; 2016;80:777–84.
27. Cava F, Hidalgo A, Berenguer J. Thermus thermophilus as biological model. Extremophiles. Springer; 2009. p. 213–31.

28. Orlygsson J, Kristjansson JK. The family Hydrogenophilaceae. The Prokaryotes: Alphaproteobacteria and Betaproteobacteria. Springer-Verlag Berlin Heidelberg; 2014. p. 859–68.
29. Ryan MP, Adley CC. Ralstonia spp.: Emerging global opportunistic pathogens. European Journal of Clinical Microbiology and Infectious Diseases. Springer; 2014. p. 291–304.
30. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. BioMed Central Ltd.; 2014;12:1–12.
31. Basso M, Venditti C, Raponi G, Navazio AS, Alessandri F, Giombini E, et al. A case of persistent bacteraemia by Ralstonia mannitolilytica and Ralstonia pickettii in an intensive care unit. Infect Drug Resist. Dove Medical Press Ltd.; 2019;12:2391–5.
32. Kim JM, Rim JH, Kim DH, Kim H, Choi SK, Kim DY, et al. Microbiome analysis reveals that Ralstonia is responsible for decreased renal function in patients with ulcerative colitis . Clin Transl Med. Wiley; 2021;11.
33. Green HD, Bright-Thomas R, Kenna DT, Turton JF, Woodford N, Jones AM. Ralstonia infection in cystic fibrosis. Epidemiol Infect. Cambridge University Press; 2017;145:2864–72.
34. Bjerre RD, Hugerth LW, Boulund F, Seifert M, Johansen JD, Engstrand L. Effects of sampling strategy and DNA extraction on human skin microbiome investigations. Sci Rep. Nature Research; 2019;9:1–11.
35. Gosiewski T, Ludwig-Galezowska AH, Huminska K, Sroka-Oleksiak A, Radkowski P, Salamon D, et al. Comprehensive detection and identification of bacterial DNA in the blood of patients with sepsis and healthy volunteers using next-generation sequencing method - the observation of DNAemia. European Journal of Clinical Microbiology & Infectious Diseases. Springer Verlag; 2017;36:329–36.
36. Païssé S, Valle C, Servant F, Courtney M, Burcelin R, Amar J, et al. Comprehensive description of blood microbiome from healthy donors assessed by 16S targeted metagenomic sequencing. Transfusion (Paris). 2016;56:1138–47.
37. Palleroni NJ. Pseudomonas . Bergey’s Manual of Systematics of Archaea and Bacteria. Wiley; 2015. p. 1–1.
38. Weyrich LS, Farrer AG, Eisenhofer R, Arriola LA, Young J, Selway CA, et al. Laboratory contamination over time during low-biomass sample analysis. Mol Ecol Resour. Blackwell Publishing Ltd; 2019;19:982–96.
39. Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, et al. Analysis of the lung microbiome in the “healthy” smoker and in COPD. PLoS One. Public Library of Science; 2011;6:16384.
40. Nakatsuji T, Chiang HI, Jiang SB, Nagarajan H, Zengler K, Gallo RL. The microbiome extends to subepidermal compartments of normal skin. Nat Commun. Nature Publishing Group; 2013;4:1–8.
41. Panaiotov S, Filevski G, Equestre M, Nikolova E, Kalfin R. Cultural Isolation and Characteristics of the Blood Microbiome of Healthy Individuals. Adv Microbiol. Scientific Research Publishing, Inc.; 2018;08:406–21.

42. Traykova D, Schneider B, Chojkier M, Buck M. Blood microbiome quantity and the hyperdynamic circulation in decompensated cirrhotic patients. PLoS One. Public Library of Science; 2017;12:e0169310.
43. Shah NB, Nigwekar SU, Kalim S, Lelouvier B, Servant F, Dalal M, et al. The Gut and Blood Microbiome in IgA Nephropathy and Healthy Controls. Kidney360. American Society of Nephrology; 2021;2:1261–74.
44. Fitzpatrick CR, Lu-Irving P, Copeland J, Guttman DS, Wang PW, Baltrus DA, et al. Chloroplast sequence variation and the efficacy of peptide nucleic acids for blocking host amplification in plant microbiome studies. Microbiome. BioMed Central Ltd.; 2018;6:1–10.
45. Martinson JNV, Pinkham N V., Peters GW, Cho H, Heng J, Rauch M, et al. Rethinking gut microbiome residency and the Enterobacteriaceae in healthy human adults. ISME Journal. Nature Publishing Group; 2019;13:2306–18.
46. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, et al. The Long-Term Stability of the Human Gut Microbiota. Science (1979). American Association for the Advancement of Science; 2013;341:1237439.
47. Oh J, Byrd AL, Park M, Kong HH, Segre JA. Temporal Stability of the Human Skin Microbiome. Cell. 2016;165:854–66.
48. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. Nature Publishing Group; 2014;505:559–63.

Figures

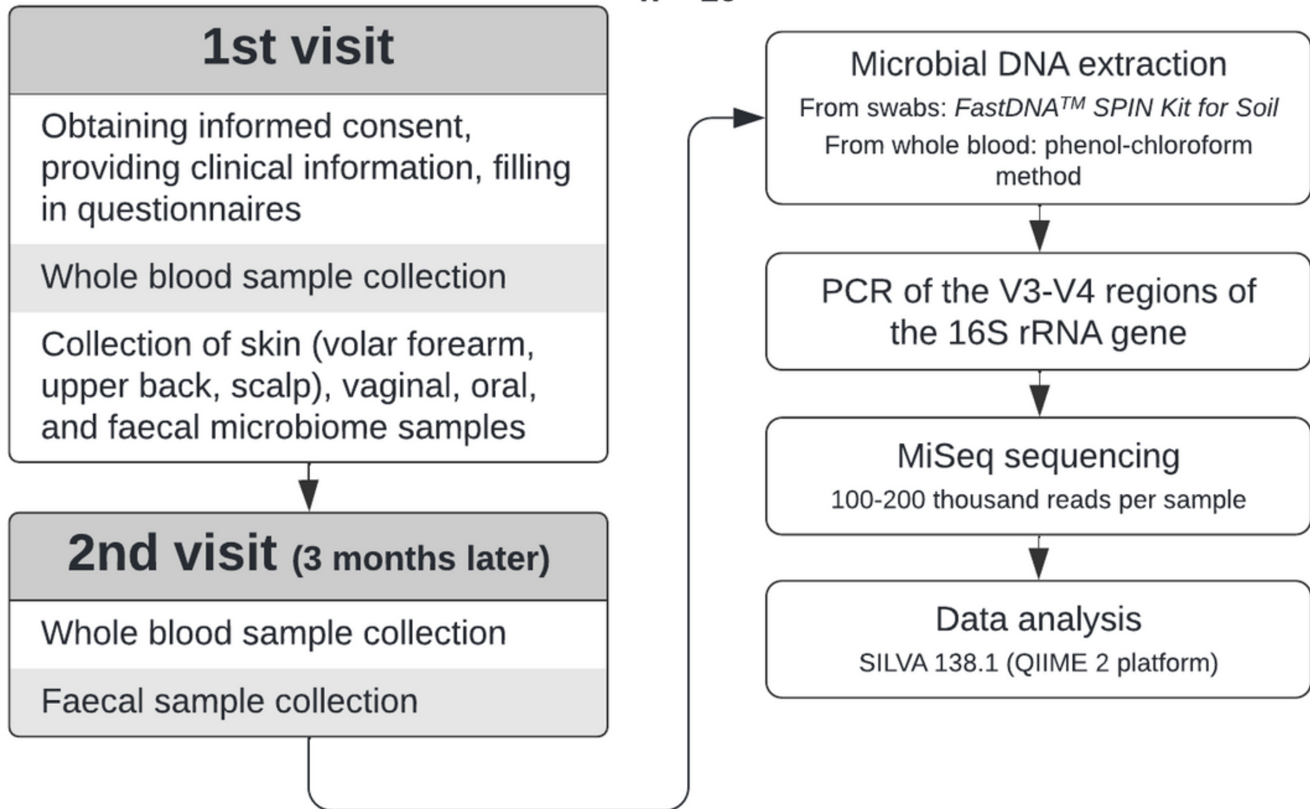

Figure 1

Study design.

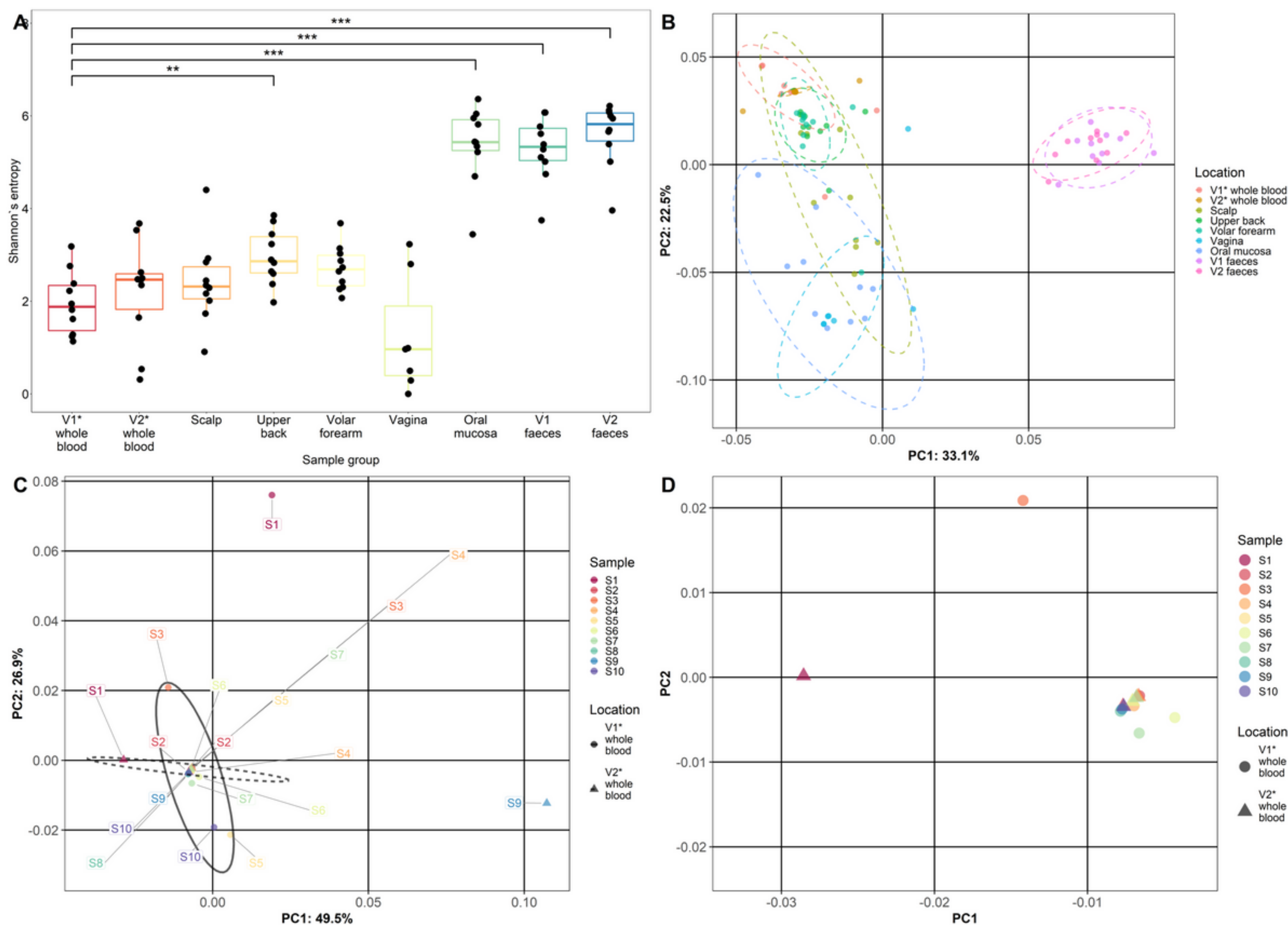

Figure 2

Alpha and beta diversity of microbiome profiles. A) Boxplots representing alpha diversity Shannon entropy index value distributions for each location. Boxplots depict median value and interquartile ranges of data in each group. Dots beyond the bounds of the whiskers represent outliers. B) Representation of beta diversity for all true samples with PCoA-weighted UniFrac distance ordination plot. Ellipses represent multivariate normal distribution (CI = 0.95) for each location. C) Representation of beta diversity only for the first and second visit blood samples with PCoA-weighted UniFrac distance ordination plot. Ellipses represent multivariate normal distribution (CI = 0.95) for each included location. D) A close-up of first and second visit blood sample PCoA-weighted UniFrac distance ordination plot (Fig. 2. C). V1: first visit, V2: second visit.

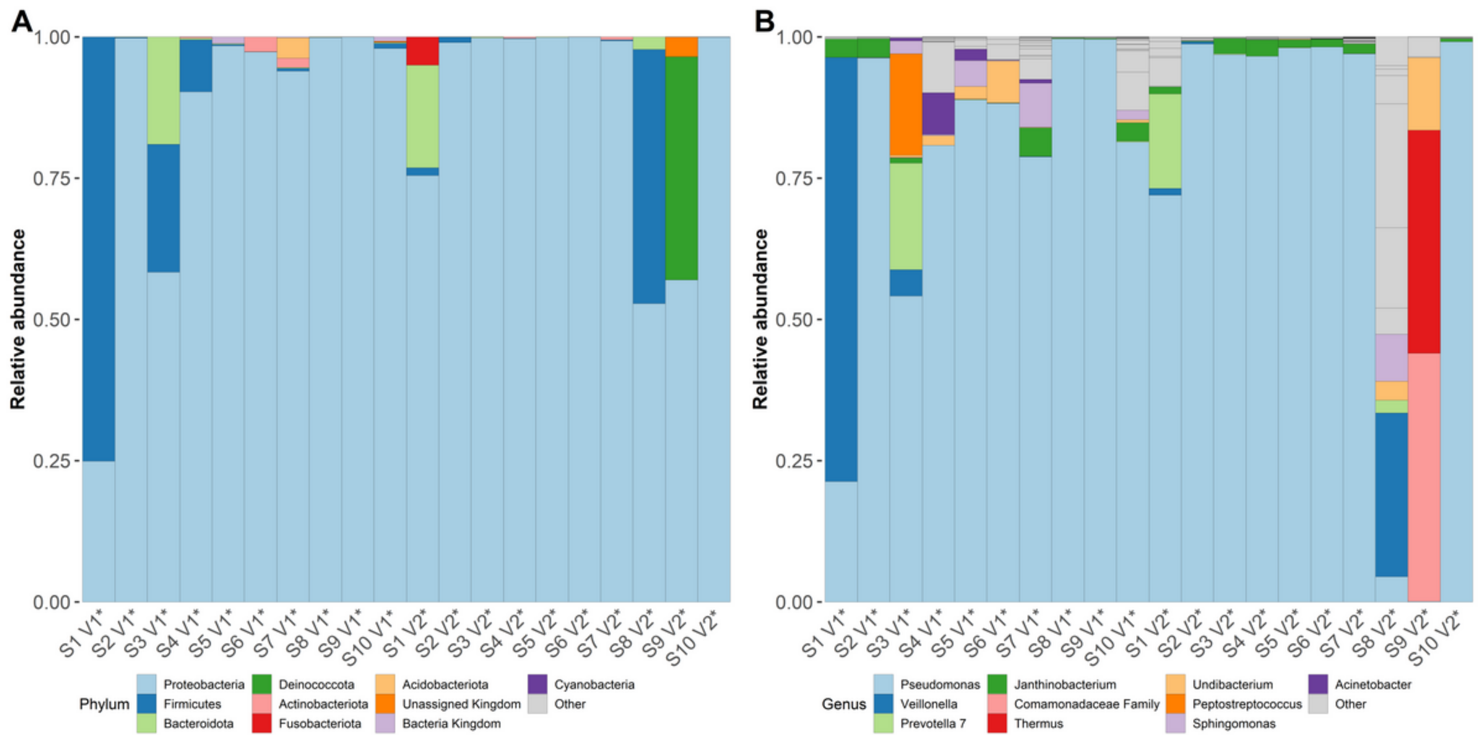

Figure 3

Taxa plot depicting relative abundances of most abundant taxa in all blood microbiome samples. A) Phyla, B) genus level. Each participant is assigned a unique identification code. V1: first visit, V2: second visit.

Supplementary Files

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

- AdditionalFile1.xlsx
- AdditionalFile2.docx