Check for updates

Preliminary landscape of *Candidatus* Saccharibacteria in the human microbiome

Sabrina Naud[1†], Camille Valles[1†], Abdourahim Abdillah[2], Linda Abou Chacra[2], Fatima Zouina Mekhalif[1], Ahmad Ibrahim[1], Aurelia Caputo[1], Jean-Pierre Baudoin[1], Frédérique Gouriet[1], Fadi Bittar[1], Jean-Christophe Lagier[1], Stéphane Ranque[2], Florence Fenollar[2], Maryam Tidjani Alou[1] and Didier Raoult[1*]

[1]Aix-Marseille Université, IRD, APHM, MEPHI, IHU Méditerranée Infection, Marseille, France, [2]Aix Marseille Université, IRD, AP-HM, SSA, VITROME, IHU Méditerranée Infection Marseille, France

Introduction: Candidate Phyla Radiation (CPR) and more specifically Candidatus Saccharibacteria (TM7) have now been established as ubiquitous members of the human oral microbiota. Additionally, CPR have been reported in the gastrointestinal and urogenital tracts. However, the exploration of new human niches has been limited to date.

Methods: In this study, we performed a prospective and retrospective screening of TM7 in human samples using standard PCR, real-time PCR, scanning electron microscopy (SEM) and shotgun metagenomics.

Results: Using Real-time PCR and standard PCR, oral samples presented the highest TM7 prevalence followed by fecal samples, breast milk samples, vaginal samples and urine samples. Surprisingly, TM7 were also detected in infectious samples, namely cardiac valves and blood cultures at a low prevalence (under 3%). Moreover, we observed CPR-like structures using SEM in all sample types except cardiac valves. The reconstruction of TM7 genomes in oral and fecal samples from shotgun metagenomics reads further confirmed their high prevalence in some samples.

Conclusion: This study confirmed, through their detection in multiple human samples, that TM7 are human commensals that can also be found in clinical settings. Their detection in clinical samples warrants further studies to explore their role in a pathological setting.

1 Introduction

The evolution of culture-independent methods has been a defining factor in the study of the human microbiome. These technological advances have led to an unprecedented understanding of the human microbiome. Moreover, these studies have also highlighted a large proportion of unidentified sequences due to the significant number of uncultured microorganisms (Lok, 2015). The exploration of this so-called dark matter has led to the creation of additional divisions within the prokaryotic domains, namely Candidate Phyla Radiation (CPR) within the bacterial domain (Hugenholtz et al., 1998; Rinke et al., 2013) and Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanohaloarchaeota, and Nanoarchaeota (DPANN) within the archaeal domain (Rinke et al., 2013). In fact, Rheims et al. described the first sequence of the 16S rRNA gene of TM7 (later renamed *Candidatus* Saccharibacteria) in 1996 (Rheims et al., 1996). This was confirmed by subsequent phylogenetic studies showing the existence of 74 divisions or putative divisions, including the TM7 division (Hugenholtz et al., 1998; Castelle et al., 2018), and the reconstruction of genomes from metagenomes in the oral microbiota (Naud et al., 2022). CPR are widely spread in the environment and are found in aquatic environments such as freshwater lakes, seawater, groundwater as well as sediments, soils and plant rhizosphere (Ji et al., 2022). For instance, *Candidatus* Parcubacteria were detected in freshwater lakes (Linz et al., 2017) whereas *Ca.* Saccharibacteria and *Candidatus* Woesebacteria were found in ocean water (Tully et al., 2018). CPR are also widely spread in the human microbiome as they are found in several niches including the oral cavity, the gastrointestinal, urogenital, and respiratory tracts as well as the skin (Naud et al., 2022). Moreover, CPR are commensals of the oral microbiota with *Ca.* Saccharibacteria and *Ca.* Absconditabacteria as the most represented phyla. Interestingly, the phylum *Ca.* Saccharibacteria is the most represented in the human microbiome and has been detected in the gastrointestinal and urogenital tracts as well as the skin microbiota (Naud et al., 2022).

In this study, we specifically investigated the presence of *Ca.* Saccharibacteria in human samples through a retrospective and prospective study using molecular biology methods complemented with scanning electron microscopy (SEM). The objective was to describe the repertoire of *Ca.* Saccharibacteria in different human niches in physiological and clinical settings.

2 Materials and methods

2.1 Ethical approval

This study primarily used anonymized samples that were not specifically obtained for this context but rather were clinical samples remaining from diagnostic screenings. The patients were informed of the possible use of their samples for research purposes and retained their right to deny approval at any point. According to the French Jardé Law (Loi n° 2012–300 du 5 mars 2012 and Décret n° 2016–1537 du 16 Novembre 2016 published in the Journal Officiel de la République Française), as this noninvasive study did not involve the specific collection of samples or the use of medical/personal data from patients, neither institutional ethical approval nor individual patient consent was required. This general approach was validated by the ethical committee of the Méditerranée Infection Institute under agreement number n° 2019-002. The breast milk samples used in this study were collected as part of another study aiming to investigate the microbial diversity of these samples. This study was validated by the national ethical committee of Senegal under approval number SEN16/45 (Sarr et al., 2021). The participants gave informed and signed consent for this study. Both studies were conducted according to the guidelines of the Declaration of Helsinki.

2.2 Screening of clinical samples using molecular techniques

Ca. Saccharibacteria are detected in samples using molecular methods, namely standard PCR (Sizova et al., 2015; Naud et al., 2022), real-time PCR (RT-PCR (Ibrahim et al., 2021; Naud et al., 2022)) as well as shotgun sequencing (Baker, 2022; Naud et al., 2022). Moreover, as *Ca.* Saccharibacteria cells have a small size and are uncultivable in axenic conditions, imaging methods such as electron microscopy and FISH are also used for their detection (Naud et al., 2022). In this study, each sample underwent two types of analyses, molecular biological and imaging (Figure 1). Molecular analyses included standard PCR, real-time PCR and shotgun sequencing whereas imaging was achieved through SEM.

2.2.1 DNA extraction
2.2.1.1 Biological fluids (oral, milk, urine, vaginal and blood samples)

DNA from clinical specimens was extracted using the EZ1 advanced XL biorobot (Qiagen, Courtaboeuf, France) with the EZ1 DNA tissue kit and the "DNA bacteria" extraction program. Two hundred microliters of each sample were used, and DNA elution was performed in a 100 µL volume.

2.2.1.2 Solid specimens (fecal samples and cardiac valves)

The extraction of bacterial DNA from human fecal samples was performed using mechanical and chemical lysis as previously described (Guindo et al., 2020). For cardiac valve samples, DNA extraction was carried out as follows: first, we added 200 µL of buffer G2 (EZ1 DNA Tissue Kit) to the sample (10-40 mg) along with 10 µL of proteinase K (EZ1 DNA Tissue Kit). The sample was incubated at 56°C with an agitation of 350 rpm for 2 h. The digested sample was transferred to a new tube containing a tip of glass powder for mechanical lysis using Fastprep (MP Biomedicals, Instrument FastPrep-24 5G) for 40 seconds at maximum speed (6.5 m/s) before performing a ten minute-incubation at 100°C. After centrifugation (30 seconds at 13000 rpm), 200 µL of the supernatant was collected. Automated extraction using the EZ1
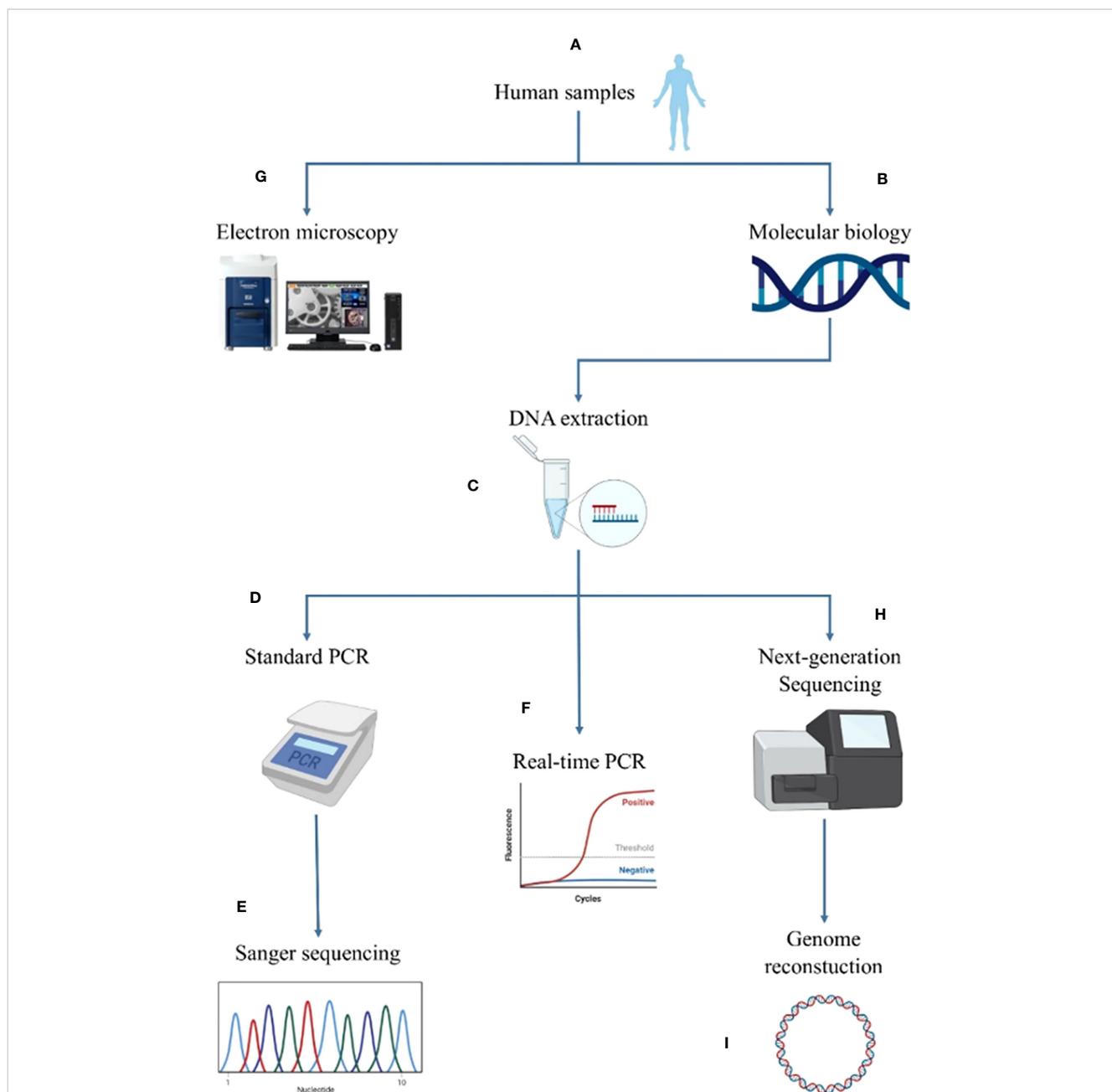
FIGURE 1

Summary of the screening methodology for *Candidatus* Saccharibacteria in the different human sites. The human samples **(A)** were first analyzed by molecular biological methods **(B)**. For this purpose, DNA was extracted **(C)** and analyzed using standard PCR **(D)** complemented by Sanger sequencing **(E)** and RT–PCR **(F)**. Imaging was performed for positive samples using tabletop electron microscopy **(G)**. Next-generation sequencing was performed for positive samples using molecular biology and electron microscopy **(H)** to reconstruct a new genome of *Candidatus* Saccharibacteria **(I)**.

biorobot and the DNA Tissue kit (Qiagen) was then performed as described above with a final elution volume of 200 µL.

2.2.2 TM7 molecular detection

TM7 molecular detection was achieved using standard PCR as well as real-time PCR (RT–PCR). Real-time PCR was performed using the primers and probe as previously described (Table S1) (Ibrahim et al., 2021). Briefly, 5 µL of DNA was mixed with 10 µL of MasterMix (Thermo Fisher, Illkirch-Graffenstaden, France), 0.5 µL

of uracil DNA glycosylase (UDG), 3 µL of DNase/RNase-free ultrapure distilled water and 0.5 µL each of the forward primer, reverse primer and probe. The amplification and hybridization reactions were performed using a Light Cycler 480 with the Roche standard program. Standard PCR was performed using the M7580F and 1177R primer pair targeting the 16S ribosomal ribonucleic acid (rRNA) gene as previously described (Sizova et al., 2015). Sanger sequencing was also performed on positive samples as previously described (Drancourt et al., 2000) using the

same primers used for standard PCR. The raw sequences were trimmed and assembled using ChromasPro software version 1.7 (Technelysium Pty Ltd., Tewantin, Australia). The obtained sequences were matched with the GenBank database using the BLASTN program (https://blast.ncbi.nlm.nih.gov/Blast.cgi, last accessed January 2023). The resulting sequences were then used to construct a phylogenetic tree using the software MEGA7 (Kumar et al., 2016) with the maximum likelihood method (Tamura 3 parameter model/method) and 500 bootstrap replicates. Alignment was conducted using Muscle v3.8.31 with default parameters.

2.2.3 Next generation sequencing

Genomic DNA was sequenced with the paired-end strategy and was barcoded to be mixed with other genomic projects prepared with the Nextera XT DNA Sample Prep Kit (Illumina Inc., San Diego, CA, USA). To prepare the paired-end library, a dilution was performed to obtain 1 ng of each sample as input. The "tagmentation" step fragmented and tagged the DNA. Then, a limited cycle PCR amplification (18 cycles) completed the attachment of the tag adapters and introduced dual-index barcodes. After purification with AMPure XP beads (Beckman Coulter Inc, Fullerton, CA, USA), the libraries were then normalized using two different methods depending on the method of sequencing. For sequencing using a MiSeq sequencer (Illumina), libraries were normalized on specific beads according to the Nextera XT protocol (Illumina) and pooled together. Automated cluster generation and paired-end sequencing with dual index reads were performed in a single 39-hour run in 2x250 bp with a MiSeq reagent Kit (V2-500 cycles) (Illumina).

2.2.4 Bioinformatic analysis

Reads were adjusted using Trimmomatic version 0.36.6 (Bolger et al., 2014) and assembled using the SPAdes software version 3.13.0 with default parameters (Bankevich et al., 2012). All contigs with a minimum length of 400 bp were conserved. BLASTn of these contigs versus nr was performed prior to the reconstruction of the genome into one single scaffold using the CONTIGuator tool (Galardini et al., 2011). Additionally, mapping of the obtained reads was performed using CLC Genomics Workbench (Liu and Di, 2020) against a reference genome of *Ca.* Saccharibacteria, *Ca.*

Nanosynbacter lyticus strain TM7x (GenBank accession number CP007496.1), using default settings with the length set at 0.9 and the similarity fraction set at 0.8. Moreover, all the reads obtained from 16S amplicon sequencing at the genomics platform of the University Hospital Méditerranée Infection Institute were analyzed using the MetaGX pipeline with the Silva 9.0 and culturomics databases (Diakite et al., 2019). MetaGX is an in-lab pipeline developed with the start-up XeGen (http://xegen.eu/) based on QIIME (Caporaso et al., 2010) using BLAST (Altschul et al., 1990) for taxonomic assignment and SILVA (Quast et al., 2013; Yilmaz et al., 2014) as a reference dictionary. For taxonomic assignments, we only included OTUs consisting of at least 20 reads. The OTUs were then searched against each database using BLASTN. The best match (≥ 97% identity and 100% coverage) was retrieved for each OTU from the reference database, and taxonomy was assigned up to the species level.

The database consisted of 3,788 samples at the time our analysis was conducted. Nonhuman samples (including animal, environmental, insect and unidentified samples) were excluded from this study. A total of 3,550 human samples were ultimately included in our analysis.

2.3 Sample screening using scanning electron microscopy

All samples except cardiac valves were fixed in a 2.5% glutaraldehyde fixative solution for at least one hour. Using the cytospin instrument, the sample was spotted on the slide and contrasted using a 1% aqueous phosphotungstic acid (PTA) solution (pH 7.0) for two minutes. The slide was then air-dried and observed using electron microscopy as previously described. Electron micrographs were acquired using a Hitachi TM4000 Plus benchtop scanning electron microscope (Hitachi, Tokyo, Japan). Cardiac valve samples were prepared with a one-hour fixation with 2.5% glutaraldehyde solution before cutting with a razor blade. The samples were rinsed twice, first with 0.1 M Caco and then with distilled water for 1 minute each. Dehydration of the cardiac valves was performed under agitation in five successive two-minute ethanol baths (20%, 50%, 70%, 90%, and 100%). We then performed two 5-min successive incubations of the valve sample, first in a bath consisting of 100% ethanol (1 V) and 100% hexamethyldisilazane (HMDS) (2 V), followed by a second bath consisting of 100% HMDS solely. The valve was dried under a laminar flow hood for 1 h and a cross-section was deposited on a glass slide. The observation was performed using a Hitachi SU5000 scanning electron microscope (Hitachi, Tokyo, Japan). In this study, we considered *Ca.* Saccharibacteria as a bacterial symbiont or free structure with widths ranging from 100 nm to 500 nm, as previously described (He et al., 2015; Cross et al., 2019; Ibrahim et al., 2021). This identification was presumptive and based on the current state of knowledge on CPR. CPR were identified based on SEM micrograph of a coculture of *Schaalia odontolytica* with *Ca.* Saccharibacteria (Figure S1) were obtained in a separate study by our team.

2.4 Statistical analysis

Statistical analyses were performed using R (Chan, 2018), GraphPad Prism 8 and OpenEpi (https://www.openepi.com/Menu/OE_Menu.htm). Normality and variances were determined using R (Shapiro test and Bartlett test, respectively). Quantitative comparisons were conducted using the Mann–Whitney test for two variables and the Kruskal–Wallis test for three variables or more.

3 Results

3.1 Prevalence of CPR in human samples from the MetaGX database

The 16S amplicon datasets from the MetaGX database were analyzed to assess the prevalence and abundance of CPR within the

TABLE 1 Relative abundance of CPR phyla in different human niches.

Samples	Candidatus Saccharibacteria	Candidatus Atribacteria	Candidatus Parcubacteria	Candidatus Microgenomates	Candidatus Dojkabacteria	Candidatus Gracilibacteria	Candidatus Katanobacteria
Oral cavity	2.92E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Sputum	2.85E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Respiratory tract	4.98E-03	7.24E-06	1.10E-05	7.08E-06	2.55E-06	1.58E-05	0.00E+00
Gastro-intestinal tract	3.84E-03	1.30E-05	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Urinary tract	3.64E-03	9.19E-05	1.47E-04	1.12E-04	2.70E-05	0.00E+00	5.61E-05
Human breast milk	2.70E-03	1.25E-04	8.42E-06	0.00E+00	0.00E+00	2.48E-06	0.00E+00

This table is presented as a heatmap.

human microbiome (Diakite et al., 2019). Seven CPR phyla, namely, *Ca.* Saccharibacteria, *Ca.* Atribacteria, *Ca.* Parcubacteria, *Ca.* Microgenomates, *Ca.* Dojkabacteria, *Ca.* Gracilibacteria and *Ca.* Katanobacteria were detected in these datasets (Table 1). *Ca.*

Saccharibacteria was the most represented phylum in all the analyzed sample types (oral cavity, sputum, gastrointestinal tract, respiratory tract, urinary tract, and human breast milk). A total of 28.9% (1,026/3,550) of the samples presented reads assigned to the CPR and the phylum *Ca.* Saccharibacteria was the most represented. One hundred percent (9/9) of the sputum samples, 94.7% (18/19) of the oral samples, 37% (210/568) of the respiratory
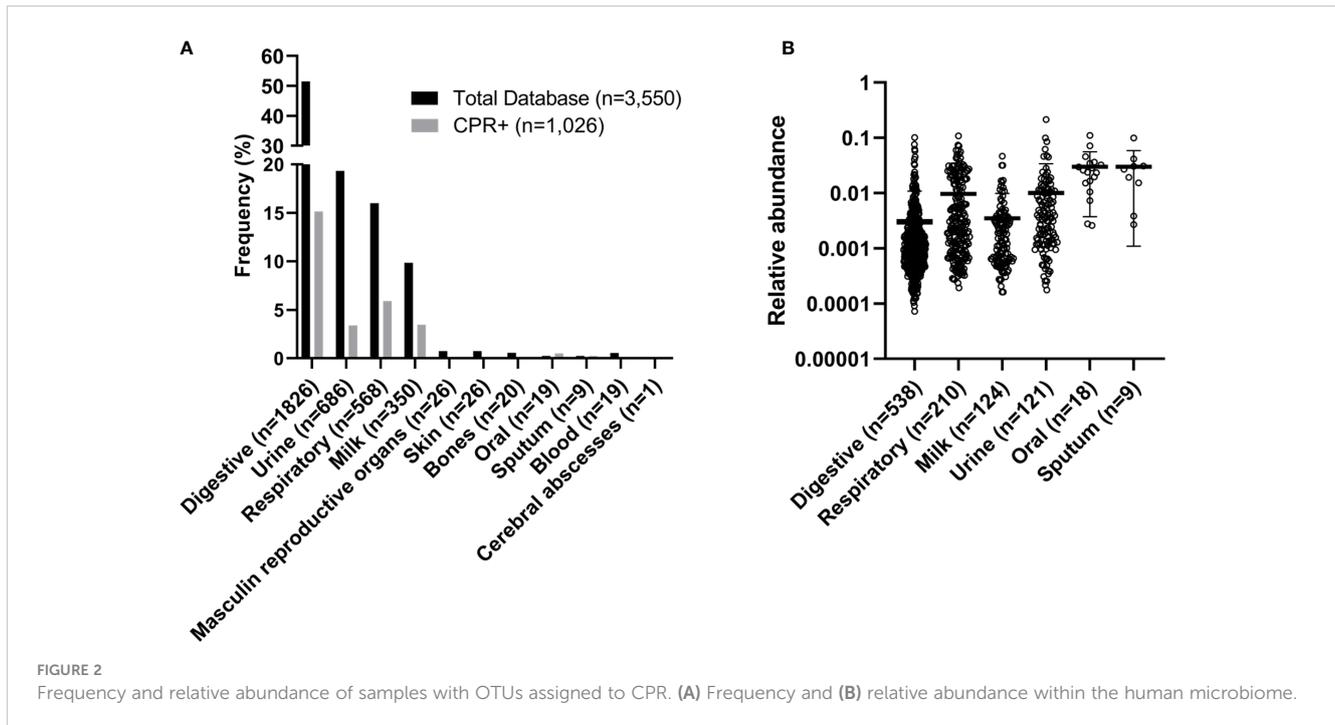
samples, 35.4% (124/350) of the breast milk samples, 29.5% (538/1,826) of fecal samples and 17.6% (121/686) of the urine samples presented OTUs assigned to *Ca.* Saccharibacteria (Table 2). A few reads assigned to *Ca.*

Saccharibacteria were found in skin, bone, and male genitalia samples (15.4% (4/26), 2% (1/50), 3.8% (1/26) samples, respectively). In addition, no reads of *Ca.* Saccharibacteria were found in human blood and cerebral abscess samples (Figure 2A). The oral microbiota (oral and sputum samples) exhibited a higher prevalence of CPR than the other studied niches (p-value<0.0001, Kruskal-Wallis test) (Figure 2B, Table S2).

TABLE 2 Summary table of the methodologies used to screen CPR in this study.

Samples	Standard PCR	Sanger Sequencing	RT–PCR	Microscopy	Mapping (Total metagenomic)	16S amplicon datasets	Positive correlation
Oral cavity	89.2% (58/65)		96.9% (95/98)		Proof of concept	94.7% (18/19)	
Gastrointestinal tract	65.8% (50/76)		52.6% (40/76)		Proof of concept	29.5% (538/1826)	
Urine	8.8% (14/160)		0.6% (1/160)			17.6% (121/686)	
Human breast milk	2.7% (5/182)		21.4% (39/182)			35.4% (124/350)	
Respiratory tract						37% (210/568)	
Sputum						100% (9/9)	
Vagina	3.1% (9/288)		6.6% (19/288)			0% (0/1)	
Blood	0.8% (63/7405)					0% (0/20)	
Cardiac valves	2.5% (17/670)		1.2% (8/693)				
Skin						15.4% (4/26)	
Bones						2% (1/50)	
Male genital organs						3.8% (1/26)	
Cerebral abscesses						0% (0/1)	

Green, methodology used with positive results; gray, methodology not used; and red, methodology used with negative results.

FIGURE 2
Frequency and relative abundance of samples with OTUs assigned to CPR. **(A)** Frequency and **(B)** relative abundance within the human microbiome.

3.2 Retrospective screening of *Candidatus* Saccharibacteria in clinical samples

After analysis of our 16S amplicon database, we performed a retrospective study and screened for the presence of *Ca.* Saccharibacteria in different types of clinical samples. The screening was performed using standard PCR and real-time PCR for the initial assessment, which was then confirmed by observation of *Ca.* Saccharibacteria using scanning electron microscopy and Sanger sequencing (Figure 3, Table S3) of the PCR products of positive samples (Table 2). Finally, we were able to reconstruct two CPR bacterial genomes from the mapping of reads obtained from shotgun metagenomics of stool and oral cavity samples (Table 2).

3.2.1 Human oral samples

Oral samples (dental plaque) exhibited the highest prevalence of *Ca.* Saccharibacteria. They were detected in 97% (95/98) of positive samples using RT–PCR and 89.2% (58/65) of positive samples using standard PCR (Figure 4A). Ct values ranged from 16.93 to 31.90 with an average Ct of 23.41 (Figure 4A). Additionally, Sanger sequencing of the PCR amplicons of approximately 600 bp showed homology with the 16S rRNA of *Ca.* Saccharibacteria (Figure 3, Table S3). Structures compatible with those of CPR (cocci-like shapes, free or associated with bacteria with a diameter under 500 nm) according to the literature (He et al., 2015; Bor et al., 2016; Ibrahim et al., 2021; Naud et al., 2022) were observed in human oral specimens (Figure 4A) using SEM in positive samples. The mapping-based assembly of reads obtained from the shotgun sequencing of a positive oral sample (Ct=19.04) allowed the reconstruction of a genome with 81% coverage of the reference genome (GenBank accession number CP007496.1). This genomic sequence consisted of one scaffold with a size of 0.7 Mbp and a GC%
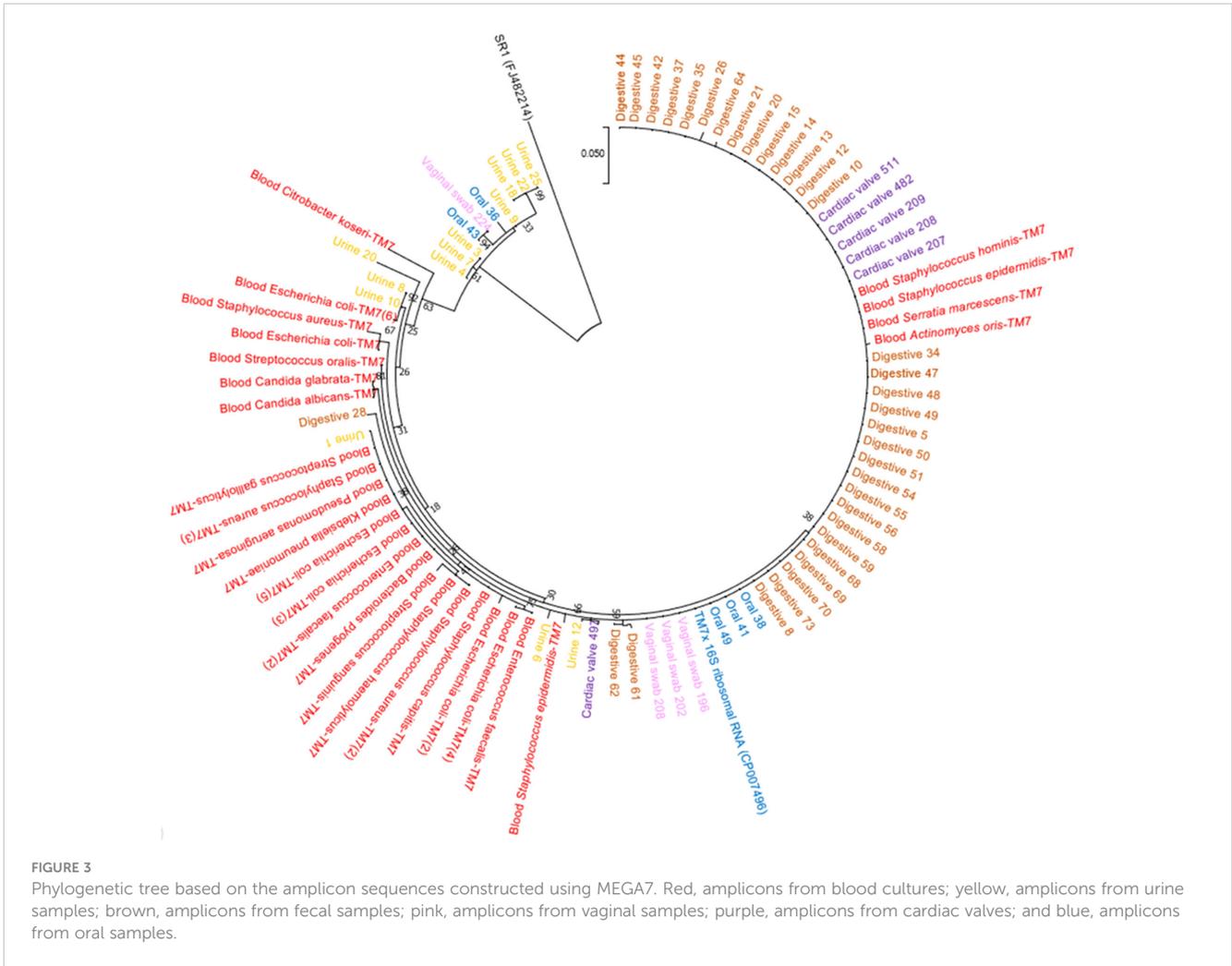
of 43.42. The 16S rRNA gene (GenBank accession number OX335640) shared a similarity of 100% with that of *Ca.* Nanosynbacter sp. HMT-352 strain TM7-037 (CP089288.1), whereas the maximum digital DNA–DNA hybridization (dDDH) (Auch et al., 2010; Meier-Kolthoff et al., 2013) of 70.7% was shared with *Ca.* Minimicrobia massiliensis (GenBank accession number CADDWL010000000).

3.2.2 Human fecal samples

Fecal samples also presented a prevalence of *Ca.* Saccharibacteria and accounted for 52.6% (40/76) of the positive samples based on the RT–PCR results, with a mean Ct of 28.55 [24.53-33.33] (Figure 4B). Additionally, 65.8% (50/76) of fecal samples were positive for *Ca.* Saccharibacteria using standard PCR (Figure 4B), and the corresponding sequences matched *Ca.* Saccharibacteria (Figure 3, Table S3). This was confirmed with the SEM micrographs showing structures consistent with these minimicrobes in positive fecal samples (Figure 4B). Moreover, two genomes (GenBank accession numbers CADDWL010000000 and CAJGBL010000000) were reconstructed from the mapping of reads generated using shotgun sequencing of two fecal samples (Ct 24.65 and Ct 27.55) against *Ca.* Nanosynbacter lyticus TM7x (GenBank accession number CP007496.1). These genomes consisted of 91 and 197 contigs, resulting in a size of 0.9 Mbp and 0.7 Mbp with a 43.2% and 43.7% GC content, respectively. These two genomic sequences shared a dDDH value of 56.8 [54.0-59.5].

3.2.3 Urogenital microbiota

Ca. Saccharibacteria were also detected in the urogenital tract. Urine samples showed a prevalence of 0.6% (1/160, Figure 4C) using RT–PCR and 8.8% (14/160) using standard PCR. A higher

FIGURE 3
Phylogenetic tree based on the amplicon sequences constructed using MEGA7. Red, amplicons from blood cultures; yellow, amplicons from urine samples; brown, amplicons from fecal samples; pink, amplicons from vaginal samples; purple, amplicons from cardiac valves; and blue, amplicons from oral samples.

prevalence was obtained in vaginal samples using RT–PCR (6.6% (19/288)), with Ct values ranging from 22.35 to 34.88 (Figure 4D), and 3.1% (9/288) for standard PCR. These results were confirmed with Sanger sequencing of the PCR products (Figure 3, Table S3) and SEM observation of structures consistent with those of CPR in positive samples (Figures 4C, D).

3.2.4 Human breast milk

Human breast milk was an unexpected niche for CPR. These samples showed a non-negligible prevalence for CPR with 21.4% (39/182) of breast milk samples positive using RT–PCR (33.00 [28.94; 34.92]) (Figure 4E), whereas a lower prevalence was obtained using standard PCR, with 2.7% (5/182) of positive breast milk samples (Figure 4E). Moreover, symbiotic structures were observed in positive breast milk samples using SEM (Figure 4E).

3.2.5 Human clinical samples (blood and cardiac valves)

Interestingly, *Ca.* Saccharibacteria were also detected in clinical samples at a low frequency. A large screening of 7,405 positive blood culture bottles from febrile patients showed that 0.8% (63/7,405) were positive using standard PCR targeting *Ca.*

Saccharibacteria (Figure 5A). These data were confirmed by Sanger sequencing of the PCR products (Figure 3, Table S3). The most common bacterial species associated with *Ca.* Saccharibacteria were *Escherichia coli* (15%), *Klebsiella pneumoniae* (11%), *Staphylococcus aureus* (11%), *Staphylococcus epidermidis* (10%), *Staphylococcus capitis* (10%), and *Enterococcus faecalis* (6%). Strikingly, we also detected *Ca.* Saccharibacteria in the blood culture of a patient infected by the yeast *Candida albicans* (3%). These results were confirmed using SEM that revealed structures consistent with CPR in positive blood cultures (Figure 5A). Cardiac valve samples showed a prevalence of 1.2% (8/693) using RT–PCR, including 75% (6/8) of control valves (valves not affected by infectious endocarditis) and 25% (2/8) of valves from patients with infective endocarditis caused by *Serratia marcescens* and *Streptococcus mitis* (Figure 5B). Standard PCR screening detected *Ca.* Saccharibacteria in 2.5% (17/670) of cardiac valve samples (Figure 5B). However, their presence could not be confirmed using SEM (Figure 5B).
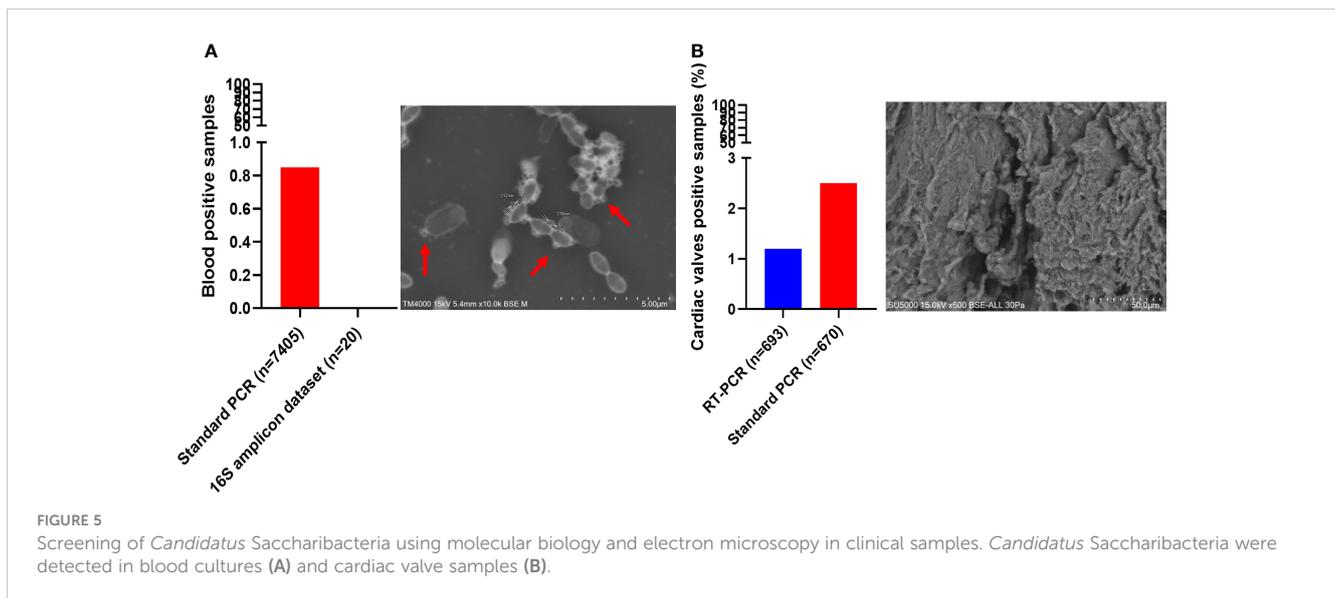
4 Discussion

Until recently, CPR were overlooked within the bacterial domain and therefore in the human microbiome. Their recent
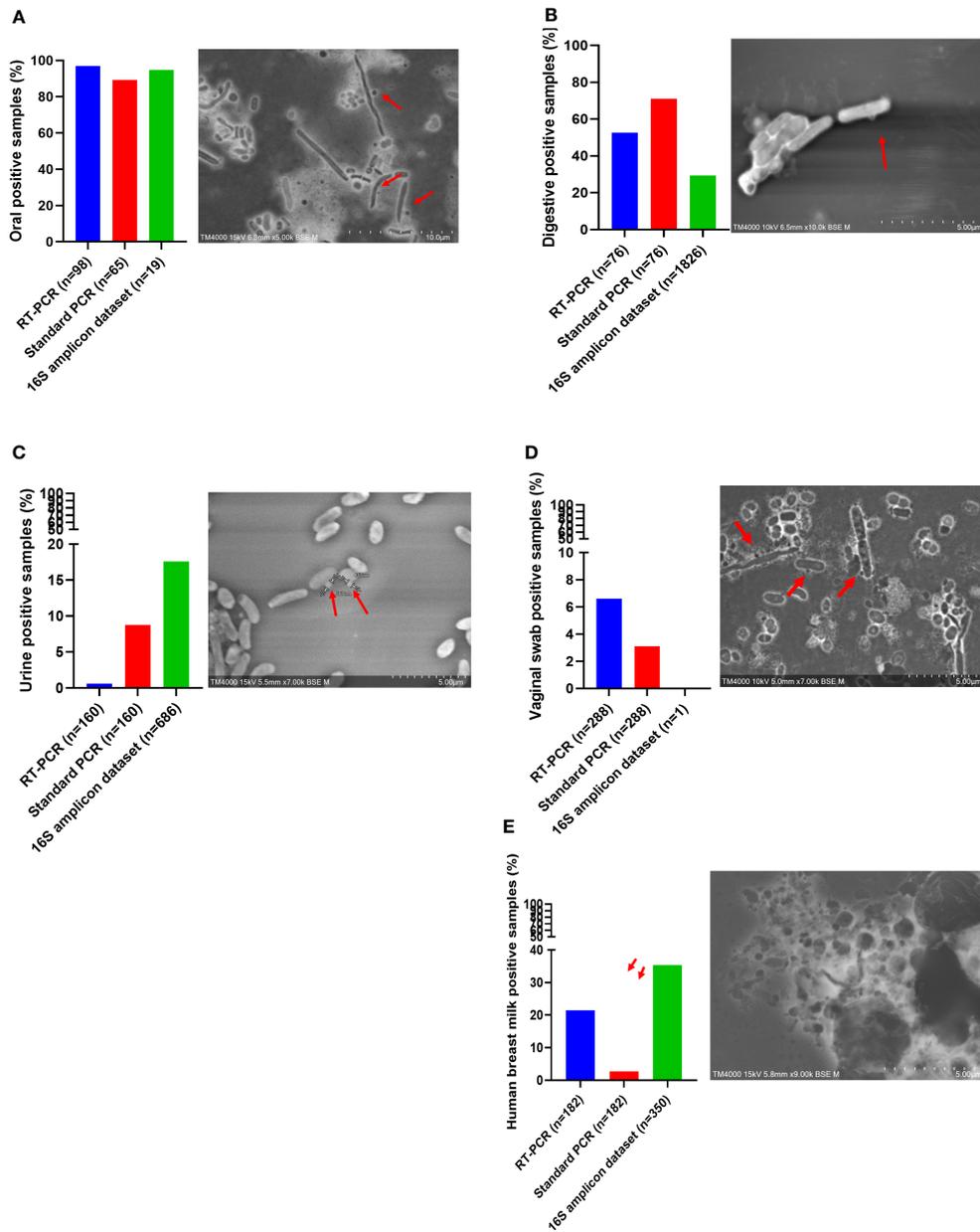
FIGURE 4
Screening of *Candidatus* Saccharibacteria using molecular biology and electron microscopy in anatomical sites. *Candidatus* Saccharibacteria were detected in oral samples **(A)**, fecal samples **(B)**, human breast milk samples **(C)**, urine samples **(D)**, and vaginal swab samples **(E)**.

discovery has led to their increased detection in 16S amplicon sequencing datasets as well as those of metagenomics, thus highlighting their non-negligible presence in the oral cavity. Here, we have carried out a preliminary study exploring the prevalence of CPR, and more specifically *Ca.* Saccharibacteria in various human sample types through a retrospective analysis of CPR within our in-lab database of 3,550 16S amplicon datasets as well as prospective analysis of readily available human samples. Thus, we highlighted the ubiquitous presence of CPR, specifically that of *Ca.* Saccharibacteria, in the human microbiome. We demonstrated the quasi-ubiquitous presence of *Ca.* Saccharibacteria in the oral cavity as previously described (Bor et al., 2016; Bor et al., 2019; Cross et al., 2019; Bor et al., 2020; Naud et al., 2022). *Ca.*

Saccharibacteria were also detected in other human niches, including the gastrointestinal tract, urogenital tract, and breast milk, as well as in clinical samples, including blood samples and heart valves. The detection of CPR through molecular methods was confirmed by visualization via electron microscopy, which revealed structures under 500 nm generally attached to bacterial hosts. It is noteworthy that better detection rates were achieved using RT–PCR targeting 23S rRNA than using standard PCR targeting the 16S rRNA gene. The RT–PCR system presented a high specificity due to its design based on all complete *Ca.* Saccharibacteria genomes available at the time of the design (Ibrahim et al., 2021), in contrast to the standard PCR system, which presented a higher sensitivity as it was designed based on the 16S rRNA gene

FIGURE 5
Screening of *Candidatus* Saccharibacteria using molecular biology and electron microscopy in clinical samples. *Candidatus* Saccharibacteria were detected in blood cultures **(A)** and cardiac valve samples **(B)**.

(Takenaka et al., 2018). Despite the discrepancy between the two methods, both highlighted the oral microbiota as exhibiting the highest prevalence of *Ca.* Saccharibacteria as described in the literature (Heller et al., 2016; Oliveira et al., 2016; Bor et al., 2020; Murugkar et al., 2020), followed by the gastrointestinal tract, breast milk, vaginal tract and urine. To our knowledge, this study represents the largest screening of CPR in humans. We provide a more complete picture of the repartition of *Ca.* Saccharibacteria within the human microbiome through the addition of breast milk as a previously unknown niche. Moreover, *Ca.* Saccharibacteria were also detected in pathogenic settings, specifically in the blood of febrile patients and cardiac valves of patients suffering from infectious endocarditis. This is the first reported detection in cardiac valves whereas *Ca.* Parcubacteria, another CPR phylum, has previously been reported within human blood (Kowarsky et al., 2017; Chen et al., 2020). This study warrants further exploration of CPR prevalence and relative abundance, including phyla other than *Ca.* Saccharibacteria, in larger sample sizes and additional niches such as the skin and respiratory tract, in which CPR have been previously detected (Zhou et al., 2013; McLean et al., 2020; Rueca et al., 2021). Furthermore, case–control studies could be conducted to specifically assess the impact of CPR on dysbiosis.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethical committee of the Méditerranée Infection Institute and the National ethical committee of Sénégal. Written informed consent was not provided because this study primarily used anonymized samples that were not specifically obtained for this context but rather were clinical samples remaining from diagnostic screenings. The patients were informed of the possible use of their samples for research purposes and retained their right to deny approval at any point. According to the French Jardé Law (Loi n° 2012–300 du 5 mars 2012 and Décret n° 2016–1537 du 16 Novembre 2016 published in the Journal Officiel de la République Française), as this noninvasive study did not involve the specific collection of samples or the use of medical/personal data from patients, neither institutional ethical approval nor individual patient consent was required. This general approach was validated by the ethical committee of the Méditerranée Infection Institute under agreement number no 2019-002. The breast milk samples used in this study were collected as part of another study aiming to investigate the microbial diversity of these samples. This study was validated by the national ethical committee of Senegal under approval number SEN16/45 (Sarr et al., 2021). The participants gave informed and signed consent for this study. Both studies were conducted according to the guidelines of the Declaration of Helsinki.

Author contributions

Conceptualization, MT and DR; Data curation, AC; Formal analysis, SN, CV, AA, LA, FM and AI; Funding acquisition, DR; Investigation, SN, CV, AA, LA, FM, AI and J-PB; Methodology, MT and DR; Resources, FG, FB and FF; Supervision, J-CL, SR, MT and DR; Validation, J-CL, SR, FF, MT and DR; Visualization, SN and CV; Writing – original draft, SN, CV, MT and DR; Writing – review & editing, MT and DR.

Funding

Acknowledgments

Conflict of interest

DR was a consultant in microbiology for the Hitachi High-Tech Corporation until March 2021.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be misconstrued as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2023.1195679/full#supplementary-material

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Auch, A. F., von Jan, M., Klenk, H.-P., and Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci.* 2, 117–134. doi: 10.4056/sigs.531120

Baker, J. L. (2022). Using Nanopore Sequencing to Obtain Complete Bacterial Genomes from Saliva Samples. *mSystems* 7 (5), e0049122. doi: 10.1128/msystems.00491-22

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bor, B., Bedree, J. K., Shi, W., McLean, J. S., and He, X. (2019). Saccharibacteria (TM7) in the Human Oral Microbiome. *J. Dent. Res.* 98, 500–509. doi: 10.1177/0022034519831671

Bor, B., Collins, A. J., Murugkar, P. P., Balasubramanian, S., To, T. T., Hendrickson, E. L., et al. (2020). Insights Obtained by Culturing Saccharibacteria With Their Bacterial Hosts. *J. Dent. Res.* 99, 685–694. doi: 10.1177/0022034520905792

Bor, B., Poweleit, N., Bois, J. S., Cen, L., Bedree, J. K., Zhou, Z. H., et al. (2016). Phenotypic and Physiological Characterization of the Epibiotic Interaction Between TM7x and Its Basibiont Actinomyces. *Microb. Ecol.* 71, 243–255. doi: 10.1007/s00248-015-0711-7

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Castelle, C. J., Brown, C. T., Anantharaman, K., Probst, A. J., Huang, R. H., and Banfield, J. F. (2018). Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat. Rev. Microbiol.* 16, 629–645. doi: 10.1038/s41579-018-0076-2

Chan, B. K. C. (2018). Data Analysis Using R Programming. *Adv. Exp. Med. Biol.* 1082, 47–122. doi: 10.1007/978-3-319-93791-5_2

Chen, L.-X., Anantharaman, K., Shaiber, A., Eren, A. M., and Banfield, J. F. (2020). Accurate and complete genomes from metagenomes. *Genome Res.* 30, 315–333. doi: 10.1101/gr.258640.119

Cross, K. L., Campbell, J. H., Balachandran, M., Campbell, A. G., Cooper, S. J., Griffen, A., et al. (2019). Targeted isolation and cultivation of uncultivated bacteria by reverse genomics. *Nat. Biotechnol.* 37, 1314–1321. doi: 10.1038/s41587-019-0260-6

Diakite, A., Dubourg, G., Dione, N., Afouda, P., Bellali, S., Ngom, I. I., et al. (2019). Extensive culturomics of 8 healthy samples enhances metagenomics efficiency. *PloS One* 14, e0223543. doi: 10.1371/journal.pone.0223543

Drancourt, M., Bollet, C., Carlioz, A., Martelin, R., Gayral, J. P., and Raoult, D. (2000). 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. *J. Clin. Microbiol.* 38, 3623–3630. doi: 10.1128/JCM.38.10.3623-3630.2000

Galardini, M., Biondi, E. G., Bazzicalupo, M., and Mengoni, A. (2011). CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol. Med.* 6, 11. doi: 10.1186/1751-0473-6-11

Guindo, C. O., Davoust, B., Drancourt, M., and Grine, G. (2020). Diversity of Methanogens in Animals' Gut. *Microorganisms* 9, 13. doi: 10.3390/microorganisms9010013

He, X., McLean, J. S., Edlund, A., Yooseph, S., Hall, A. P., Liu, S.-Y., et al. (2015). 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. doi: 10.1073/pnas.1419038112

Heller, D., Helmerhorst, E. J., Gower, A. C., Siqueira, W. L., Paster, B. J., and Oppenheim, F. G. (2016). Microbial Diversity in the Early I*n Vivo*-Formed Dental Biofilm. *Appl. Environ. Microbiol.* 82, 1881–1888. doi: 10.1128/AEM.03984-15

Hugenholtz, P., Goebel, B. M., and Pace, N. R. (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol* 180, 4765–4774. doi: 10.1128/JB.180.18.4765-4774.1998

Ibrahim, A., Maatouk, M., Rajaonison, A., Zgheib, R., Haddad, G., Bou Khalil, J., et al. (2021). Adapted Protocol for Saccharibacteria Cocultivation: Two New Members Join the Club of Candidate Phyla Radiation. *Microbiol. Spectr.* 9, e0106921. doi: 10.1128/spectrum.01069-21

Ji, Y., Zhang, P., Zhou, S., Gao, P., Wang, B., and Jiang, J. (2022). Widespread but Poorly Understood Bacteria: Candidate Phyla Radiation. *Microorganisms* 10, 2232. doi: 10.3390/microorganisms10112232

Kowarsky, M., Camunas-Soler, J., Kertesz, M., De Vlaminck, I., Koh, W., Pan, W., et al. (2017). Numerous uncharacterized and highly divergent microbes which colonize humans are revealed by circulating cell-free DNA. *Proc. Natl. Acad. Sci. U.S.A.* 114, 9623–9628. doi: 10.1073/pnas.1707009114

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Linz, A. M., Crary, B. C., Shade, A., Owens, S., Gilbert, J. A., Knight, R., et al. (2017). Bacterial Community Composition and Dynamics Spanning Five Years in Freshwater Bog Lakes. *mSphere* 2, e00169–e00117. doi: 10.1128/mSphere.00169-17

Liu, C.-H., and Di, Y. P. (2020). Analysis of RNA Sequencing Data Using CLC Genomics Workbench. *Methods Mol. Biol.* 2102, 61–113. doi: 10.1007/978-1-0716-0223-2_4

Lok, C. (2015). Mining the microbial dark matter. *Nature* 522, 270–273. doi: 10.1038/522270a

McLean, J. S., Bor, B., Kerns, K. A., Liu, Q., To, T. T., Solden, L., et al. (2020). Acquisition and Adaptation of Ultra-small Parasitic Reduced Genome Bacteria to Mammalian Hosts. *Cell Rep.* 32, 107939. doi: 10.1016/j.celrep.2020.107939

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.-P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinf.* 14, 60. doi: 10.1186/1471-2105-14-60

Murugkar, P. P., Collins, A. J., Chen, T., and Dewhirst, F. E. (2020). Isolation and cultivation of candidate phyla radiation Saccharibacteria (TM7) bacteria in coculture with bacterial hosts. *J. Oral. Microbiol.* 12, 1814666. doi: 10.1080/20002297.2020.1814666

Naud, S., Ibrahim, A., Valles, C., Maatouk, M., Bittar, F., Tidjani Alou, M., et al. (2022). Candidate Phyla Radiation, an Underappreciated Division of the Human Microbiome, and Its Impact on Health and Disease. *Clin. Microbiol. Rev.* 35 (3), e0014021. doi: 10.1128/cmr.00140-21

Oliveira, R. R. D. S., Fermiano, D., Feres, M., Figueiredo, L. C., Teles, F. R. F., Soares, G. M. S., et al. (2016). Levels of Candidate Periodontal Pathogens in Subgingival Biofilm. *J. Dent. Res.* 95, 711–718. doi: 10.1177/0022034516634619

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Rheims, H., Rainey, F. A., and Stackebrandt, E. (1996). A molecular approach to search for diversity among bacteria in the environment. *J. Ind. Microbiol. Biotechnol.* 17, 159–169. doi: 10.1007/BF01574689

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J.-F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437. doi: 10.1038/nature12352

Rueca, M., Fontana, A., Bartolini, B., Piselli, P., Mazzarelli, A., Copetti, M., et al. (2021). Investigation of Nasal/Oropharyngeal Microbial Community of COVID-19 Patients by 16S rDNA Sequencing. *Int. J. Environ. Res. Public Health* 18, 2174. doi: 10.3390/ijerph18042174

Sarr, M., Tidjani Alou, M., Delerce, J., Khelaifia, S., Diagne, N., Diallo, A., et al. (2021). A Listeria monocytogenes clone in human breast milk associated with severe acute malnutrition in West Africa: A multicentric case-controlled study. *PloS Negl. Trop. Dis.* 15, e0009555. doi: 10.1371/journal.pntd.0009555

Sizova, M. V., Doerfert, S. N., Gavrish, E., and Epstein, S. S. (2015). TM7 detection in human microbiome: Are PCR primers and FISH probes specific enough? *J. Microbiol. Methods* 114, 51–53. doi: 10.1016/j.mimet.2015.05.005

Takenaka, R., Aoi, Y., Ozaki, N., Ohashi, A., and Kindaichi, T. (2018). Specificities and Efficiencies of Primers Targeting Candidatus Phylum Saccharibacteria in Activated Sludge. *Materials (Basel)* 11, E1129. doi: 10.3390/ma11071129

Tully, B. J., Graham, E. D., and Heidelberg, J. F. (2018). The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. *Sci. Data* 5, 170203. doi: 10.1038/sdata.2017.203

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014). The SILVA and "All-species Living Tree Project (LTP)". *taxonomic frameworks nucleic Acids Res.* 42, D643–D648. doi: 10.1093/nar/gkt1209

Zhou, Y., Gao, H., Mihindukulasuriya, K. A., La Rosa, P. S., Wylie, K. M., Vishnivetskaya, T., et al. (2013). Biogeography of the ecosystems of the healthy human body. *Genome Biol.* 14, R1. doi: 10.1186/gb-2013-14-1-r1