Journal of Dentistry 146 (2024) 105059

Contents lists available at ScienceDirect

Journal of Dentistry

journal homepage: www.elsevier.com/locate/jdent

Metagenomic analysis reveals ecological and functional signatures of oral phageome associated with severe early childhood caries

Xin Chen [a,e], Ting Zou [b], Qinglu Zeng [c], Yubing Chen [c], Chengfei Zhang [d], Shan Jiang [b,*], Guicong Ding [a,e,**]

[a] Shenzhen Children's Hospital of China Medical University (CMU), Shenzhen, PR China
[b] Shenzhen Stomatology Hospital (Pingshan), Southern Medical University, Shenzhen, Guangdong, PR China
[c] The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, PR China
[d] Endodontology, Restorative Dental Sciences, Faculty of Dentistry, The University of Hong Kong, Hong Kong, PR China
[e] Department of Stomatology, Shenzhen Children's Hospital, Shenzhen, PR China

ARTICLE INFO

Keywords:
Severe early childhood caries
Metagenomics
Oral microbiome
Oral virome
Salivary phageome

ABSTRACT

Objectives: Severe early childhood caries (S-ECC) is highly prevalent, affecting children's oral health. S-ECC development is closely associated with the complex oral microbial microbiome and its microorganism interactions, such as the imbalance of bacteriophages and bacteria. Till now, little is known about oral phageome on S-ECC. Therefore, this study aimed to investigate the potential role of the oral phageome in the pathogenesis of S-ECC.

Methods: Unstimulated saliva (2 mL) was collected from 20 children with and without S-ECC for metagenomics analysis. Metagenomics sequencing and bioinformatic analysis were performed to determine the two groups' phageome diversity, taxonomic and functional annotations. Statistical analysis and visualization were performed with R and SPSS Statistics software.

Results: 85.7 % of the extracted viral sequences were predicted from phages, in which most phages were classified into *Myoviridae, Siphoviridae*, and *Podoviridae*. Alpha diversity decreased, and Beta diversity increased in the S-ECC phageome compared to the healthy group. The abundance of *Podoviridae phages* increased, and the abundance of *Inoviridae, Herelleviridae,* and *Streptococcus* phages decreased in the S-ECC group. Functional annotation revealed increased annotation on glycoside hydrolases and nucleotide metabolism, decreased glycosyl transferases, carbohydrate-binding modules, and biogenic metabolism in the S-ECC phageome.

Conclusions: Metagenomic analysis revealed reduced *Streptococcus* phages and significant changes in functional annotations within the S-ECC phageome. These findings suggest a potential weakening of the regulatory influence of oral bacteria, which may indicate the development of innovative prevention and treatment strategies for S-ECC. These implications deserve further investigation and hold promise for advancing our understanding and management of S-ECC.

Clinical significance: The findings of this study indicate that oral phageomes are associated with bacterial genomes and metabolic processes, affecting the development of S-ECC. The reduced modulatory effect of the oral phageome in counteracting S-ECC's cariogenic activity suggests a new avenue for the prevention and treatment of S-ECC.

Abbreviations and acronyms: S-ECC, Severe early childhood caries; dmfs, decayed, missing, or filled surface score; vOTUs, Viral operational taxonomic units; KEGG, Kyoto Encyclopedia of genes and genomes; CAZY, Carbohydrate-active enzymes; EggNOG, Evolutionary genealogy of genes; PCoA, Principal coordinate analysis; GHs, Glycoside hydrolases; GTs, Glycosyl transferases; CBMs, Carbohydrate-binding modules; NGS, Next generation sequencing; EPS, Exopolysaccharide; PCR, polymerase chain reaction; FISH, fluorescence in situ hybridization.

* Corresponding author at: Shenzhen Stomatology Hospital (Pingshan), Southern Medical University, Shenzhen, Guangdong, PR China
** Corresponding author at: Shenzhen Children's Hospital of China Medical University (CMU), Shenzhen, China; Department of Stomatology, Shenzhen Children's Hospital, Shenzhen, PR China
E-mail addresses: kqshanj@163.com (S. Jiang), dgc_67@sina.com (G. Ding).

https://doi.org/10.1016/j.jdent.2024.105059
Received 15 December 2023; Received in revised form 5 May 2024; Accepted 7 May 2024
Available online 25 May 2024

1. Introduction

Severe early childhood caries (S-ECC) is defined as decayed, missing, or filled surfaces (dmfs), the score of which is no less than one under the age of 3, four at the age of 3, five at the age of 4, and six at the age of 5 [1]. S-ECC has become increasingly prevalent worldwide; for example, the prevalence of S-ECC increased from 9.8 % in 2013–2014 to 11.9 % in 2017–2018 in the United States [2]. S-ECC may result in dental pain, impaired mastication, malnutrition, sleeping disturbance, and distraction, impacting children's oral health-related quality of life [3,4].

Frequent microbial interactions within the oral microbiome are highly relevant to dental caries. Among the components of the oral microbiome, a robust interaction exists between bacteriophages (also known as phages) and bacteria [5]. Phages can infect bacteria, and bacteria can evolve to defend against phage infection. Cariogenic bacteria have been the focus of research on S-ECC in previous studies. In reality, numerous phages form a phage community (also called phageome) to interact directly with bacteria and firmly shape the oral microbiome [6]. Due to the duration, specificity, and efficiency of phage infection against bacteria, phages have the potential to help maintain a balance in the microbiome against pathological challenges.

The association between the oral phageome and S-ECC remains unclear. There are numerous unknown phages in the human body and the surrounding environment. At the quantity level, phages are the most prominent members of the human virome and the human microbiome, potentially composing the dominant proportion among virome and surpassing bacteria at a ratio ranging from 10 to 50 [7,8].

Due to the large number of phages, conventional cultivation methods, transmission electron microscope, polymerase chain reaction profiling, and clustered regularly interspaced short palindromic repeats profiling have limitations on studying these unknown and numerous phages. Encouragingly, metagenomics has proven effective in investigating the oral phageome and uncovering novel phages [9,10]. Metagenomics also helps advance phage therapy development for various medical conditions, including those affecting the epidermal, gastrointestinal, cardiovascular, and endocrine systems [11-16]. Thus, metagenomics is potentially an ideal resort for tentatively exploring the oral phageome in S-ECC.

Metagenomic methods facilitate the investigation of the oral phageome at both the taxonomic and functional levels [17,18]. Metagenomic analysis revealed high levels of Actinomyces phages and frequent gene-sharing connections in the phageome of adults, particularly involving Streptococcus phages [19]. Another study on adult oral phageome also reported cross-infection and an abundance of specific phages infecting Actinomyces, Corynebacterium, Lactococcus, Mycobacterium, Pseudomonas, Rhodococcus, and Streptococcus [20]. Similarly, a metagenomic analysis on children 6 to 12 years old with and without dental caries revealed differences in taxonomic and functional levels. Streptococcus phage is relevant to caries' presence, while Haemophilus phage is associated with the absence of caries. Functionally, biosynthesis pathways of urate, vitamin K2, and polyamines were observed elevated in the caries phageome, while three deiminases and lactate dehydrogenase were observed elevated in the caries-free phageome [21]. However, studies on oral phageome analysis in younger children are still limited.Therefore, applying metagenomics to investigate ecological and functional properties is critical. This study investigates the links between the oral phageome and S-ECC through metagenomic analysis of saliva samples taken from children younger than 6 years.

2. Materials and methods

2.1. Subjects and sample collection

The participating children met the following criteria: (1) aged between 3 and 5 years old; (2) willingness to comply with dental examination, sample collection, and any necessary dental treatment procedures; (3) absence of a history of systematic diseases or congenital diseases; (4) no presence of other active oral diseases; (5) provision of informed consent by the guardians.

Subsequently, all eligible children underwent a comprehensive dental examination, including a dmfs assessment conducted by trained resident dentists. Based on the dmfs scores, children were categorized into two groups: (i) a healthy group, consisting of children with a dmfs index of 0 ($n = 20$), and (ii) a S-ECC group, consisting of children with dmfs index that meet the criteria for S-ECC ($n = 20$).

Saliva samples were collected from participating children between 8:00AM and 10:00AM to minimize diurnal variations that could affect saliva composition. Prior to collection, children were instructed to abstain from eating and drinking for at least 2 h, and refrain from oral hygiene procedures for at least 1 hour to maintain a consistent oral environment [22]. To avoid any influence from medical treatments, intravenous infusions were also prohibited during this period.

Sterile saliva collection tubes, each labeled with a unique identifier, were utilized to ensure sample traceability and participant anonymity. The collection was conducted in a designated ward treatment room, ensuring optimal temperature, humidity, and a quiet environment. Children were guided to expectorate saliva into the collection tubes under the supervision of trained resident dentist. It was ensured that the volume of each saliva sample collected exceeded 2 mL to meet the minimum requirements for subsequent analyses.

Post-collection, saliva samples were promptly stored in a portable refrigerator to preserve their biochemical integrity. Within a 3-hour window, the samples were transferred to a −20 °C freezer for long-term storage, a critical step to prevent degradation and maintain sample stability. Quality control measures were rigorously implemented throughout the collection process to ensure sample quality and consistency. Samples were excluded based on the following criteria: insufficient saliva volume, contamination of the collection tube, and instances where children failed to comply with the collection procedure. Each sample was marked with a unique code corresponding to the participant's ID and the date of collection, allowing for data analysis and tracking without compromising participant privacy.

2.2. Sample preparation and metagenomic analysis

Saliva samples were mixed with 400–800 μL of saline-magnesium buffer and were filtered with 0.45 μm and 0.22 μm filters. The filtrates were treated with RNase and DNase at 37 °C for 60 min, followed by incubation at 75 °C for 10 min [23]. DNA was then extracted using a phenol-chloroform method and subjected to 1.2 % agarose gel electrophoresis to assess purity and integrity [24]. DNA concentration was measured with the Qubit® dsDNA Assay Kit in a Qubit® 3.0 Fluorometer (Life Technologies, CA, USA). Random fragmentation of DNA was conducted with the Covaris M220 Ultrasonic instrument (Covaris, Woburn, MA, USA) at a threshold of 350 bp. The DNA library was constructed with the NEB Next Ultra DNA Library PrepKit for Illumina (NEB, USA) and quantified with the Qubit® dsDNA Assay Kit in a Qubit® 3.0 Fluorometer (Life Technologies, CA, USA). Sequencing was conducted with the Illumina NovaSeq 6000 PE150 at TinyGen Biotechnology Co., Ltd. (Shanghai, China) [25]. The quality of raw sequence reads was controlled with fastp (version 0.22.0), with a lower quality limit of 20 and a length limit of 50 bp [26]. Raw reads were then mapped against GRCh38/hg38 with Bowtie 2 (version 2.3.5) [27,28]. The filtered reads were aligned and assembled into contigs with MEGAHIT (version 1.2.9) [29].

2.3. Bioinformatics analysis

The procedure of bioinformatic analysis is illustrated in Fig. 1. Viral sequences were predicted with VirFinder (version 1.1), DeepVirFinder (version 1.0), VirSorter2 (version 2.2.3), CAT (version 5.2.3), and IMG/

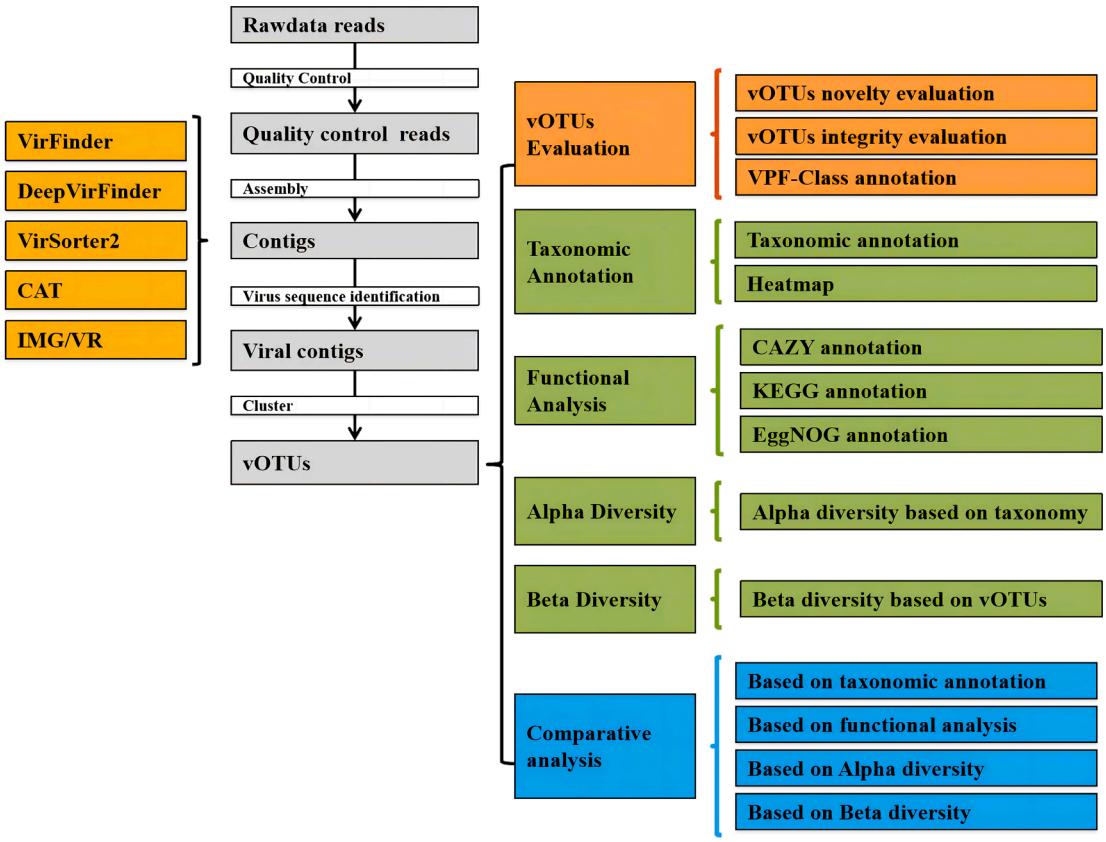

Fig. 1. Flowchart of bioinformatic analysis. The raw data went through quality control assembly to get valid contigs. Contigs were screened with VirFinder, DeepVirFinder, VirSorter2, CAT, and IMG/VR to get viral contigs. Viral contigs were clustered into vOTUs for further bioinformatic analysis. Bioinformatic analysis contains vOTUs evaluation, taxonomy annotation, Functional analysis, and diversity analysis (including Alpha and Beta diversity). The difference between the healthy group and the S-ECC group was compared based on bioinformatic data.

VR (version 5.1) [30-33,13]. Obtained sequences were clustered into viral operational taxonomic units (vOTUs) with CD-HIT (version 4.8.1) and predicted for genetic information with prodigal (version 2.6.3) [34, 35]. Reported vOTUs were selected based on a similarity of at least 90 % and coverage of at least 75 % in the comparison. Any vOTUs that did not meet these criteria were considered novel or unreported vOTUs. Taxonomic annotation was conducted by DIAMOND (version 2.0.15) against the NCBI NR database [36,37]. Functional annotations were performed against the Kyoto Encyclopedia of Genes and Genomes (KEGG), Carbohydrate-Active Enzymes (CAZY), and Evolutionary Genealogy of Genes (EggNOG) databases [38-40]. The host prediction was inferred from the IMG/VR database with VPF-Class (version 0.1.2) [41]. R software (version 3.6.1) was applied for statistical analysis and visualization, and *t*-tests (SPSS Statistics software, version 20.0) were employed to compare ecological and functional traits between the healthy group and the S-ECC group.

3. Results

3.1. Overview of the phageome

The 40 recruited children had an average age of 4.46 years. Saliva samples were collected from 20 healthy children and 20 children with S-ECC. Raw reads were processed by removing low-quality and contamination reads from nonviral resources to obtain clean and virus-like reads. The summary of reads in the optimization process is listed in Table 1. A total of 50.26 % of the contigs were ranged from 3001 bp to 4000 bp (Fig. 2). 37,273 vOTUs were obtained, constructed of 26,076 novel vOTUs (70.0 %) and 11,197 reported vOTUs (30.0 %). The longest 20 reported/novel vOTUs were illustrated in Table 2, the vOTUs were identified as dsDNA viruses (33,974 vOTUs), ssDNA viruses (103 vOTUs), retroviruses (21 vOTUs). There are 3.235 vOTUs remained unclassified.

Table 1
Optimization of sequencing data for desired viral contigs.

Sequencing data	Sequencing data statistics
Raw reads	2474,203,582
Clean reads	2405,632,064
Removed reads	2101,742,798
Virus-like reads	372,460,784
Total viral contigs	2920,390

3.2. Taxonomic annotation

The taxonomic annotation showed a high abundance of Siphoviridae sp., Podoviridae sp., and Myoviridae sp. at the family level and Minivirus, Chlorovirus, and Rahariannevirus at the species level in the entire oral phageome (Fig. 3).

3.3. Host prediction

The comparison of host prediction revealed that 85.7 % of the predicted hosts were bacteria. The phage abundance was significantly lower in the healthy group than in the S-ECC group, and phages infecting Bacillus, Mycobacterium, Streptococcus, Cellulophaga and Arthrobacter at the genus level were abundant in phageome in both the healthy group and the S-ECC group (Fig. 4). The heatmap illustrated the abundance levels of phages across the healthy group and the S-ECC group, with

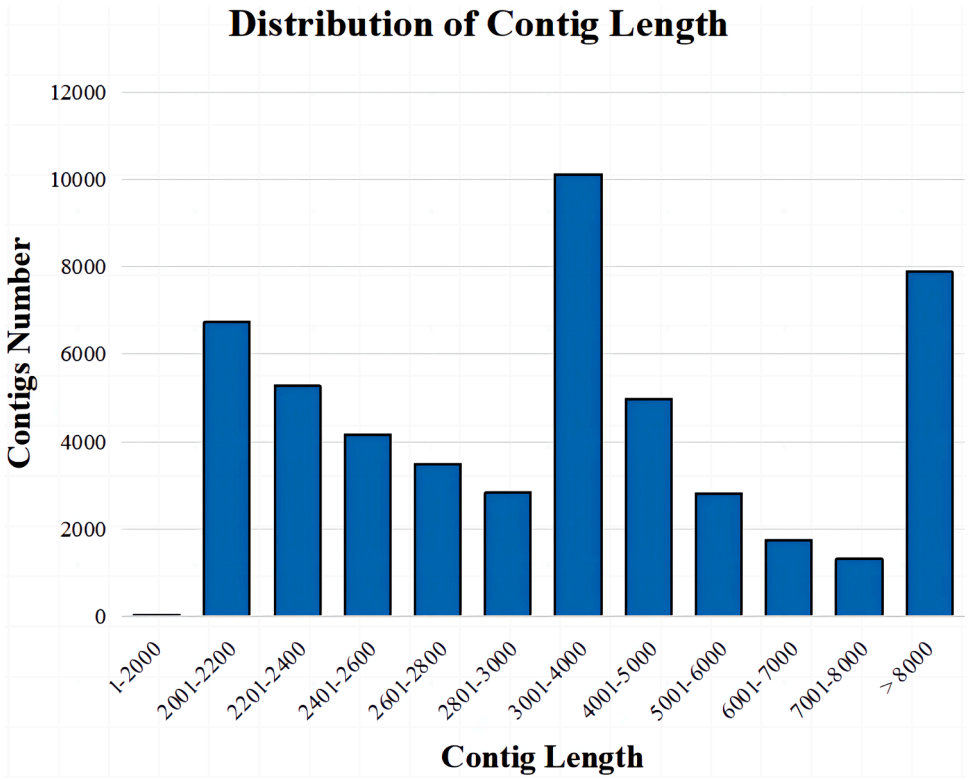

Fig. 2. Distribution of sequence length.

Table 2
The longest 20 reported vOTUs annotation with their subject ID and vOTUs length. The Subject ID refers to the sequences in the database to which the vOTUs sequences correspond.

vOTUs	Subject ID	Length
H1_C29123	IMGVR_UViG_3,300,008,638_000006|3,300,008,638|Ga0111421_100,002	392,282
H8_C129697	IMGVR_UViG_3,300,008,522_000006|3,300,008,522|Ga0111045_100,001	388,587
C3_C37668	IMGVR_UViG_3,300,008,638_000006|3,300,008,638|Ga0111421_100,002	387,948
C13_C117513	UGV-GENOME-0,380,221	381,654
H12_C119655	IMGVR_UViG_3,300,008,636_000008|3,300,008,636|Ga0111420_100,004	292,364
C16_C9730	IMGVR_UViG_3,300,008,514_000002|3,300,008,514|Ga0111023_100,008	284,814
C15_C20057	IMGVR_UViG_3,300,008,734_000001|3,300,008,734|Ga0113998_100,007	279,585
H15_C14160	IMGVR_UViG_3,300,007,208_000001|3,300,007,208|Ga0103291_100,007	279,164
H14_C4701	IMGVR_UViG_3,300,008,091_000009|3,300,008,091|Ga0105975_100,007	278,718
C7_C54240	IMGVR_UViG_3,300,008,091_000009|3,300,008,091|Ga0105975_100,007	277,608
C13_C77029	IMGVR_UViG_3,300,008,747_000008|3,300,008,747|Ga0115677_100,001	273,287
H14_C106743	UGV-GENOME-0,380,013	273,275
C1_C120126	IMGVR_UViG_3,300,008,679_000005|3,300,008,679|Ga0115430_1,000,001	269,528
H15_C96531	DTR_410,932	269,310
H18_C90306	UGV-GENOME-0,379,991	269,013
H1_C136649	UGV-GENOME-0,379,994	266,081
H12_C77777	IMGVR_UViG_3,300,008,082_000002|3,300,008,082|Ga0105968_1,000,005|2951–259,879	263,472
H12_C22198	UGV-GENOME-0,379,927	260,952
C7_C37431	IMGVR_UViG_3,300,007,980_000001|3,300,007,980|Ga0114366_1,000,001	260,355
C14_C4595	IMGVR_UViG_3,300,007,648_000001|3,300,007,648|Ga0105531_100,010	258,590

phages infecting Bacillus, Mycobacterium, Streptococcus, Cellulophaga and Arthrobacter at the genus level and Streptococcaceae, Microroccaceae, Bacillaceae, Flavobacteriaceae and Mycobacteriaceae at the family level were universally abundant in phageome in all samples (Fig. 5).

3.4. Comparison of phageome characteristics

The Alpha diversity was tested and compared at the Sobs, Chao, and ACE indexes level (Fig. 6A). The Sobs, Chao, and ACE indexes were significantly higher in the healthy group (26.85, 27.15, and 27.12, respectively) than in the S-ECC group (26.85, 27.15 and 27.12, respectively) ($p < 0.05$). The Bray-Curtis index with principal coordinate analysis (PCoA), showed higher Beta diversity in the S-ECC group (Fig. 6B). The first principal coordinates) explained 15.21 % of the variation, while the second principal coordinates) explained 12.39 % of the variation.

Taxonomic analysis indicates a higher abundance of Podoviridae and a lower abundance of Inoviridae and Herelleviridae in the S-ECC group than in the healthy group. The comparison based on taxonomic annotation revealed significantly higher abundance of Podoviridae at the family level and Podoviridae sp. at the species level in the S-ECC group than in the healthy group ($p < 0.01$) (Fig. 6C).

In a comparison of host predictions, phages infecting Streptococcus, Nitrincola, Synechoccoccus, Rhodobacter, and Klebsiella at the host genus level were observed to decrease in the S-ECC phageome, as compared with the healthy group ($p < 0.05$). The comparison based on host prediction revealed significantly lower abundance of phages infecting Streptococcus at the genus level in the S-ECC group than in the healthy group ($p < 0.01$) (Fig. 6D).

3.5. Functional annotations

The functional annotation was conducted against the KEGG, EggNOG, and CAZy databases. The annotation against KEGG predicted a

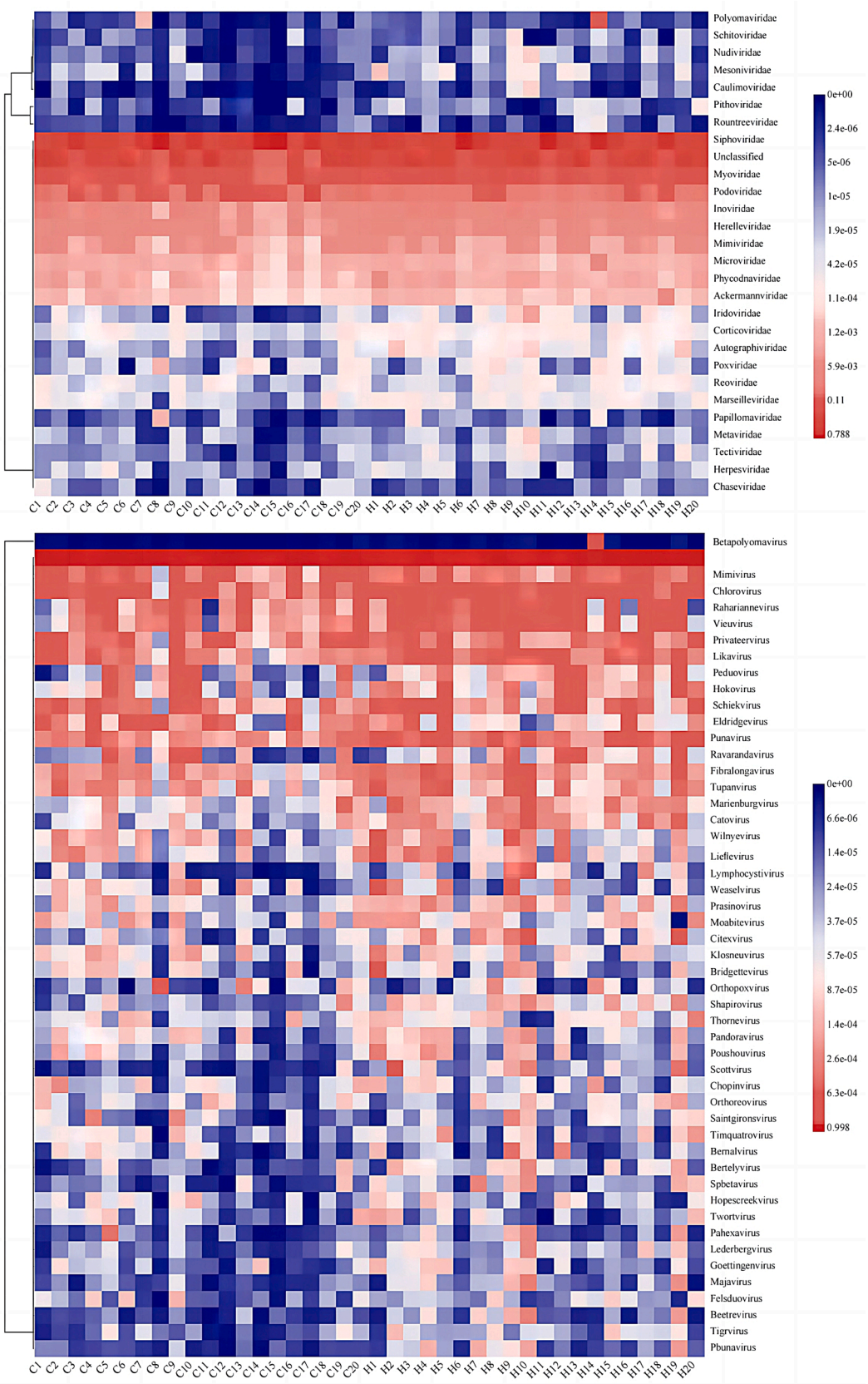

Fig. 3. Heatmap illustrating composition and taxonomic annotation in all samples at the family and species levels. The upper figure represents the taxonomic composition and annotation at the family level for all samples, while the lower figure illustrates the composition and annotation at the species level. *Siphoviridae* sp., *Podoviridae* sp., and *Myoviridae* sp. are abundant in both groups. C: S-ECC group, H: healthy group. Red cells in the heatmap indicate higher values of specific phages in each sample, and blue cells represent lower values of phages in each sample.

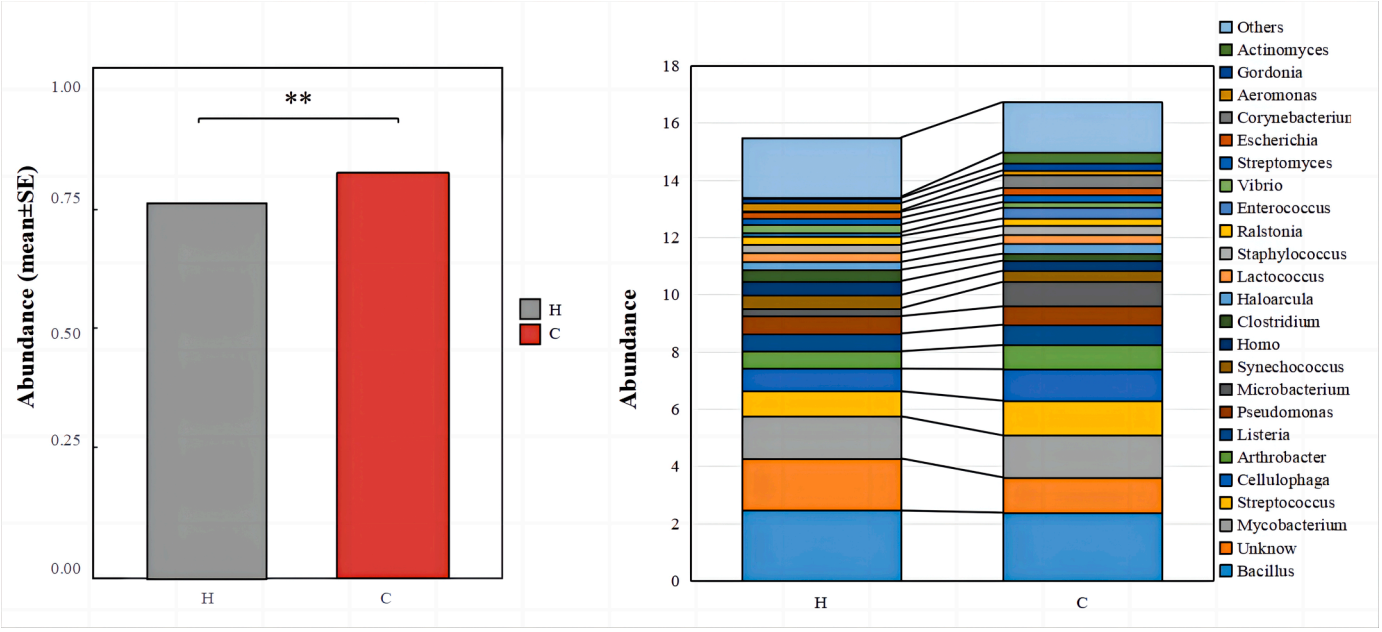

Fig. 4. Comparison and composition of vOTUs abundance with annotated bacterial hosts. The left figure represents the comparison of phages between the healthy group and the S-ECC group, while the right figure illustrates the composition of phageome at the host genus level H: healthy group, C: S-ECC group.

higher abundance of replication and repair, nucleotide metabolism, and lower carbohydrate metabolism, translation, amino acid metabolism, and energy metabolism in the S-ECC group than in the healthy group ($p < 0.05$) (Fig. 7A). The EggNOG annotation revealed an elevated abundance of nucleotide transport, nucleotide metabolism, and reduced translation, ribosomal structure, and biogenesis in the S-ECC group than in the healthy group ($p < 0.05$) (Fig. 7B). The CAZy annotation predicted elevated abundance of glycoside hydrolases (GHs) and decreased glycosyl transferases (GTs) and carbohydrate-binding modules (CBMs) in the S-ECC group than in the healthy group ($p < 0.05$) (Fig. 7C).

4. Discussion

The potential impact of the oral phageome on the bacteriome and its relation to S-ECC remains unclear, suggesting the importance of investigating oral phageome in younger children. However, challenges arise from both the low abundance of host bacteria and the viral specificity of phages, which can result in specific phages inadvertently overlooked by conventional methods. With the advancement of metagenomic technologies from polymerase chain reaction (PCR), fluorescence in-situ hybridization (FISH), DNA microarray to next-generation sequencing (NGS), the numerous and unknown microbiota can be explored more effectively than ever before [42]. Compared to other methods, NGS provides more comprehensive and extensive genetic information. NGS involves highly parallel sequencing during the sequencing and data processing steps and high-throughput instruments for large-scale, multi-channel sequencing reactions. These methodological and technological properties enable NGS to process millions to billions of DNA fragments simultaneously without isolating phages [43]. As a result, the ability to explore diverse and previously unknown microorganisms is significantly enhanced. In the present study, Illumina NovaSeq sequencing was employed to analyze the diversity and novelty of phages within the oral phageome. Illumina NovaSeq allowed for effective analysis of a wide range of microorganisms that are challenging to isolate, thus enabling a deeper and broader understanding of the oral phageome.

To the best of our knowledge, this is the first study to investigate the potential role of the phageome in the pathogenesis of S-ECC. We have identified phageome alterations associated with S-ECC, including changes in ecological diversity, taxonomic composition, and functional annotation. It was reported that a diverse phageome has been observed to exert a more rapid suppression of the host population [44]. In the present study, the Alpha diversity was lower in the S-ECC group, which aligns with the previous studies on the caries microbiome [45]. The reduced diversity may weaken its modulatory effect against bacteria. The current analysis also detected increased Beta diversity associated with S-ECC. The host prediction suggested that 85.7 % of the viral sequences originate from phages. This high abundance of phage aligns with findings from the previous study [46]. The taxonomic annotation further revealed the prevalence of Siphoviridae, Podoviridae, and Myoviridae, which is consistent with adult caries studies [47].

Due to the robust interaction between phages and bacteria, previous studies on oral bacteriomes pave the way for exploring the oral phageome. Streptococcus is the most implicated bacteria with proven cariogenicity [48]. Coincidentally, previous metagenomics studies found Streptococcus phages prevailing in oral phageome, indicating their role in the oral microbiome [49].

The current analysis revealed a lower abundance of Streptococcus phages in the S-ECC group than in the healthy group. This contrasts with the previously reported finding that more Streptococcus phages were detected in the caries phageome [21]. This difference could be due to variations in the metagenomic methodology. In the present study, the decontamination procedure removed bacterial DNA, favoring the detection of lytic phages. However, in the previous study, bacterial DNA was not removed, so lysogenic phages were included in their microbiome metagenomic analysis [21]. The decreased cariogenicity of Streptococcus lytic phages may weaken their modulatory effect on bacterial activity, thus suggesting an imbalance in phages and bacteria in developing S-ECC [50]. Furthermore, the abundance of lytic Streptococcus phages in the healthy group suggests that the healthy population may be more promising for the isolation of therapeutic phages. The presumption is also supported by the successful isolation of lytic Streptococcus mutans phages M102, e10, f1, ϕAPCM01, and SMHBZ8, from healthy subjects [51-53].

The KEGG annotation results showed an enrichment of genes associated with "nucleotide metabolism" and "replication and repair", along with a depletion of genes related to "translation". Similarly, the EggNOG annotation results revealed an enrichment of genes associated with

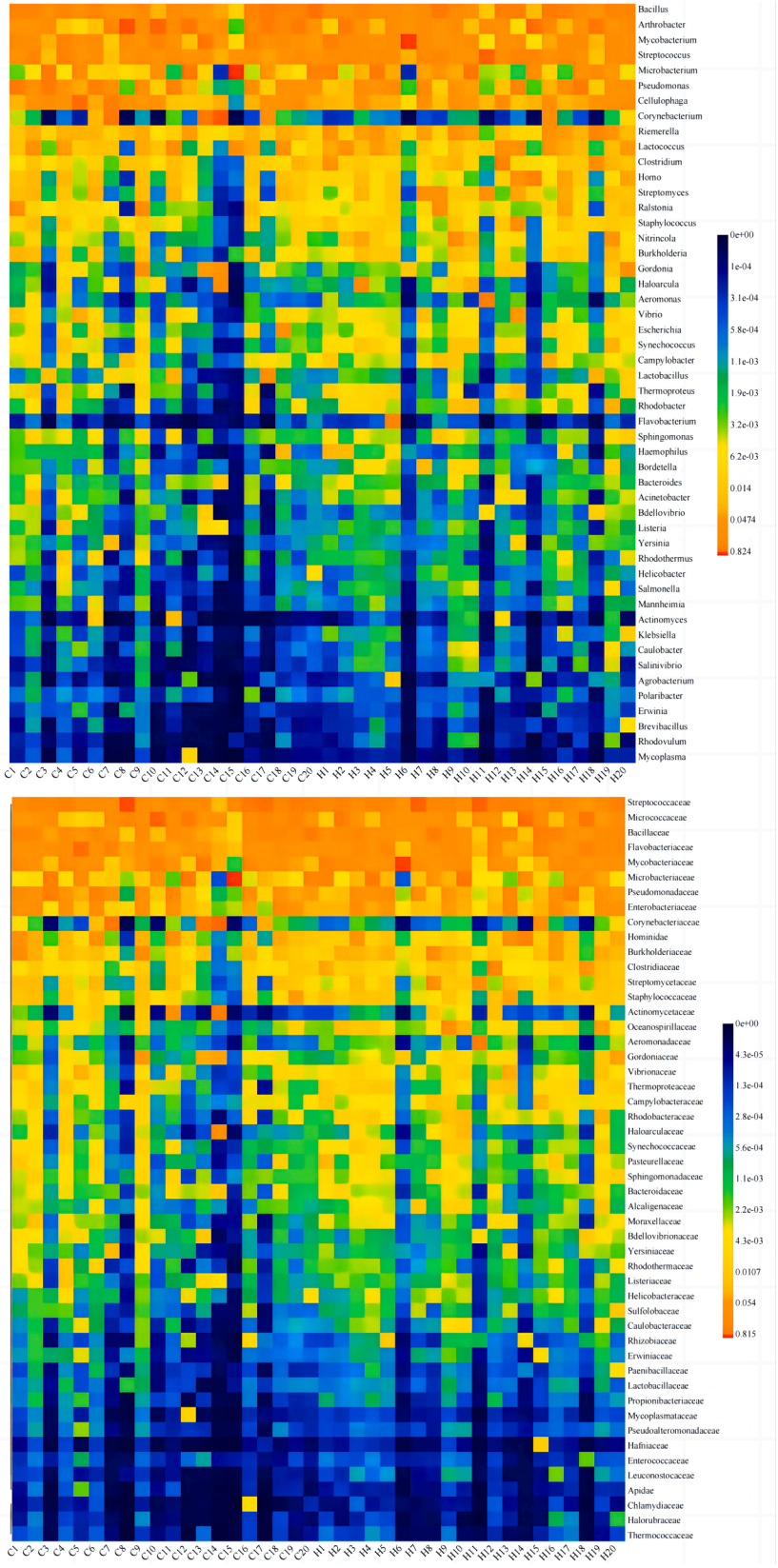

Fig. 5. Heatmap illustrating composition and host prediction in all samples at the family and species level. The left figure represents the host prediction at the host family level for all samples, while the right figure illustrates the host prediction at the host species level. Red cells in the heatmap indicate higher values of specific phages in each sample, and blue cells represent lower values of phages in each sample.

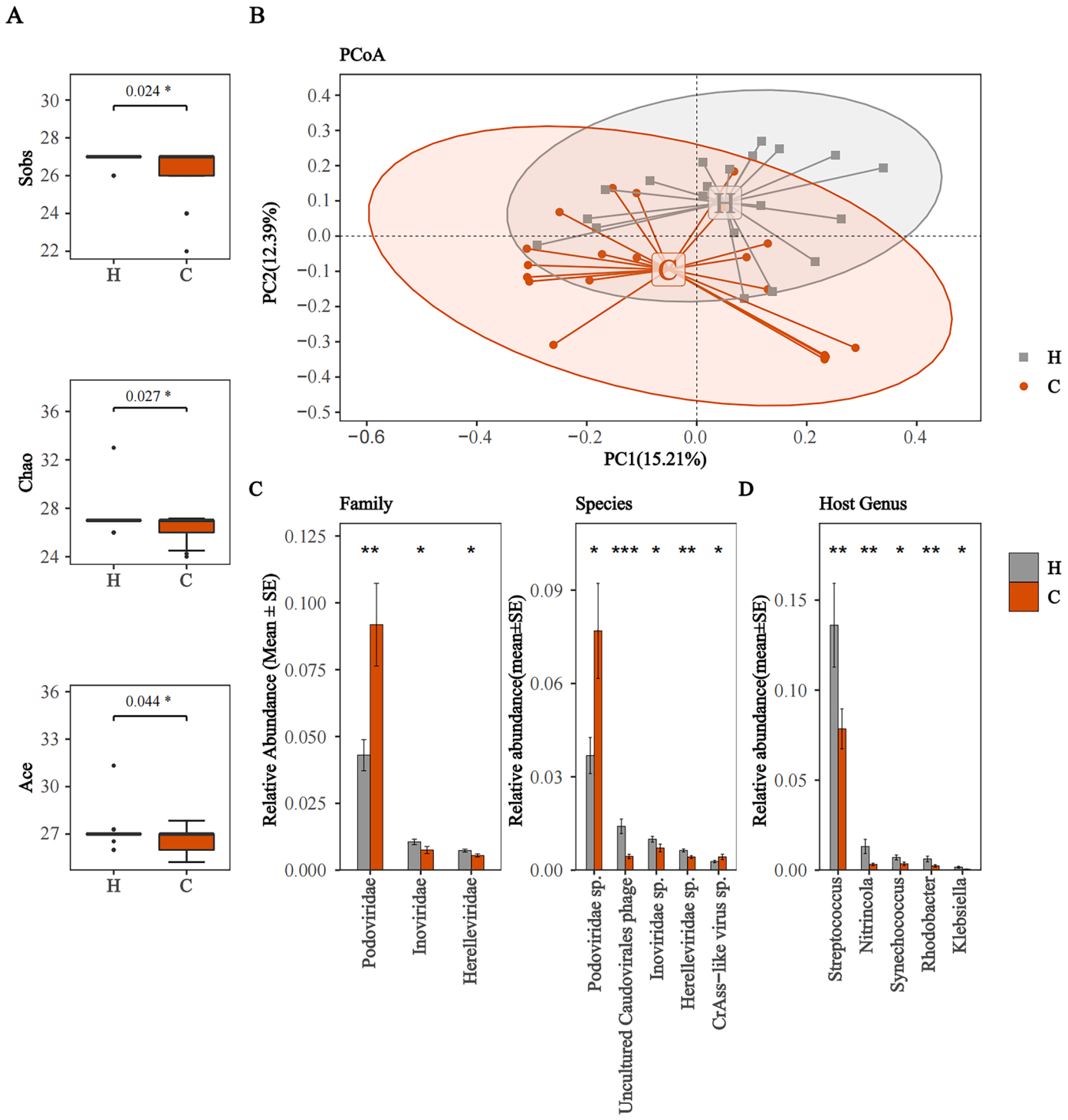

Fig. 6. Comparative analysis of composition and taxonomic annotation between the healthy group and the S-ECC group. (A). Comparative analysis of Alpha diversity (B). PCoA analysis of Beta diversity (C). Comparative analysis of the abundance of phages at the level of families and species. (d). Comparative analysis of the abundance of predicted hosts at the family and genus level. * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$, *** $0 < p < 0.001$, H: healthy group, C: S-ECC group.

"nucleotide transport and metabolism" and a depletion of genes related to "translation, ribosomal structure, and biogenesis". These findings collectively confirm that there are more genes associated with nucleotide metabolism and fewer genes associated with synthesis in the S-ECC phageome than in the healthy phageome. The nucleotide metabolism helps reduce mutation rates, which weakens phage flexibility to counteract with their host bacteria [54,55]. Moreover, the impact on synthesis could also prevent oral phageome from exerting a modulatory effect on the oral microbiome, for that synthesis is essential for phages' survival and virulence [56]. CAZy annotations revealed an enrichment of GHs and decreased GTs in the S-ECC group. GHs and GTs exert adverse effects on exopolysaccharides (EPS) [57,58]. GHs could hinder bacterial attachment and biofilm formation by breaking down EPS and polysaccharides [59]. Contrarily, GTs catalyze the synthesis of EPS [60]. The changes in functional annotations in the S-ECC group indicate that oral phageome may present modulatory effects against S-ECC development or undergo adaptive changes in response to S-ECC. Further studies are needed to substantiate this hypothesis.

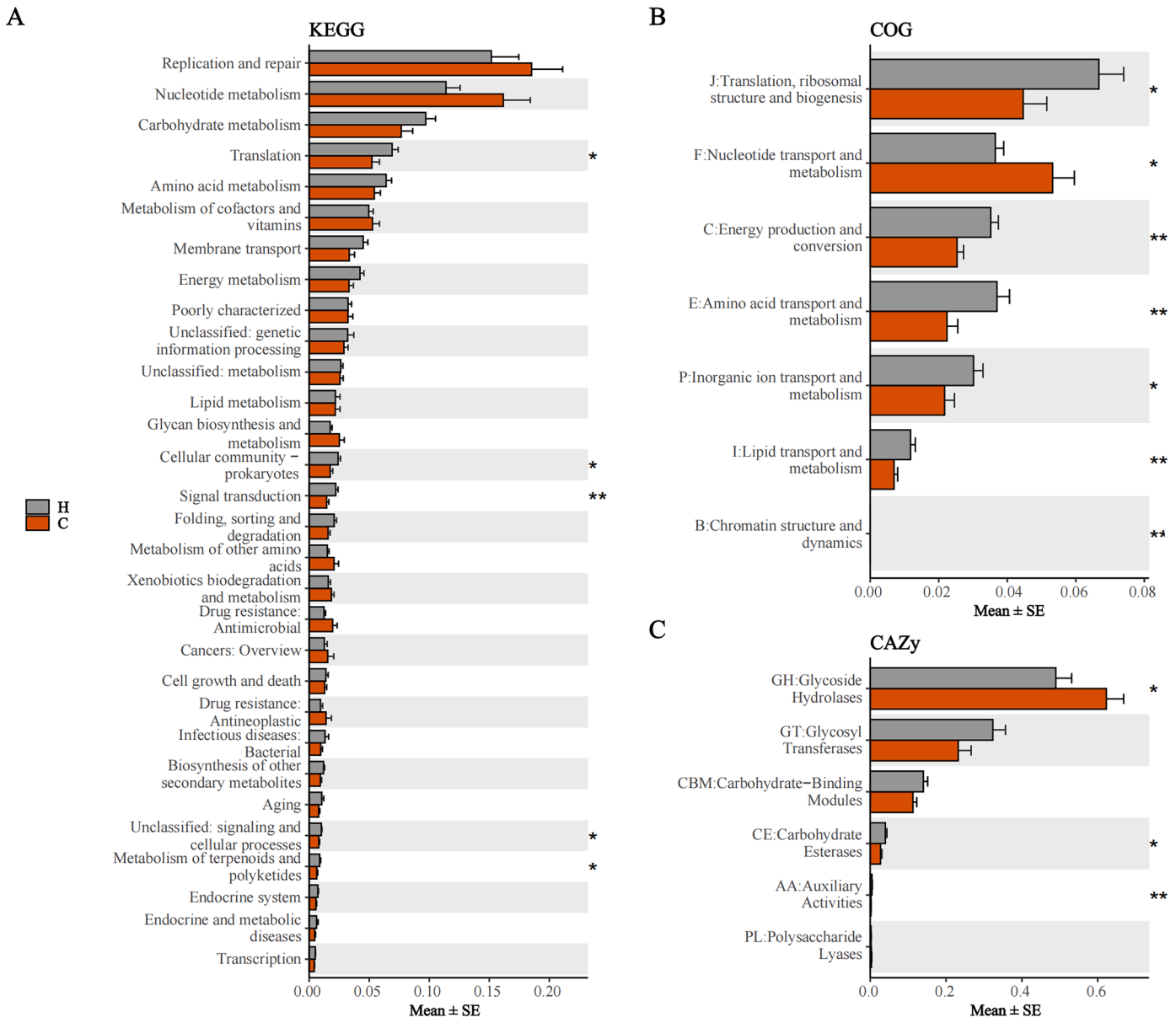

Fig. 7. Comparative analysis of functional annotation between the healthy and S-ECC groups. (A). Comparative analysis of pathways mapped against the KEGG database. (B). Comparative analysis of genetic functions mapped against the EggNOG database. (C). Comparative analysis of predicted enzyme expression mapped against the CAZy database. * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$, *** $0 < p < 0.001$, H: healthy group, C: S-ECC group.

The above-mentioned findings show that the S-ECC phageome has a reduced modulatory effect than the healthy phageome. Additionally, the inhibitive activities of the phageome against the development of S-ECC were lessened. Given the ecological and functional changes in the S-ECC phageome, the study hypothesizes that phage-bacteria interactions play a role in S-ECC development. Phages may inhibit S-ECC by infecting cariogenic bacteria, enhancing anti-cariogenic synthesis, and directly or indirectly impeding bacterial attachment.

Saliva is an ideal sample for phageome analysis in children subjects. Based on the developmental stage of children under the age of 6, saliva serves as an accessible collection method without invasion and suffering [61]. Moreover, saliva holds diagnostic potential in oral and some systemic diseases for it being individually specific and stable over time [62]. Therefore, saliva samples were applied in the present study to investigate the relationship between oral phageome and S-ECC. It is worth noting that saliva and plaque are two representative types used for omics analysis. Further, the properties of phageome from saliva and plaque are different [20]. To be specific, the salivary phageome has been observed to have a higher taxonomic abundance, whereas the plaque phageome, particularly in patients with periodontitis, exhibits a considerably lower abundance. Furthermore, the predominant phages also vary between the salivary and plaque phageome, with Yersinia phage phiR1–37 being predominant in saliva, and Lactococcus phage 1706 and Actinomyces phage AV-1 being predominant in plaque. In addition, other oral samples, such as gingival crevicular fluid, mucous membranes and mouthwashes, might have their own distinct phageome [63-66]. These insights are speculative, based on the characteristics of the disease and the sample; The actual disparity requires confirmation through microbiome analysis. Thus, an initial exploration of the salivary phageome serves as a fundamental step towards further in-depth and updated studies. Such future research efforts are critical to uncovering the microbial mechanisms underlying oral pathologies.

This study explored oral phageome in children with and without S-ECC. The demographic characteristics of the subjects, sample size, and the progression pattern of S-ECC were conservatively controlled. Future studies should consider enrolling a larger population for representative

samples. Moreover, the phageome may vary with S-ECC progression, so further studies should consider whether the S-ECC exhibits a progressive or arrested pattern. Furthermore, the relation between the phageome ecological and functional properties with the S-ECC metabolism requires more robust evidence. A combination of phageome metagenomics and bacteriome metagenomics is also needed to reach mutual corroboration on how phage-bacteria interaction functions in S-ECC development. The combination of genomics, transcriptomics, and proteomics is also essential to confirm the effect phageome play in the development of S-ECC.

5. Conclusion

In summary, our metagenomic analysis of the oral phageome in children with S-ECC has revealed a significant reduction in phages targeting Streptococcus, accompanied by significant alterations in functional annotation. These findings suggest a reduced ability to affect the composition and metabolic processes of the oral bacteria, thus highlighting the importance of the oral phageome in the development of S-ECC. The reduced modulatory effect to counteract cariogenic factors may pave the way for novel prevention and treatment strategies for S-ECC. Future research should consider integrating metagenomics targeting both the phageome and the bacteriome to understand better the phage-bacteria interaction associated with S-ECC.

CRediT authorship contribution statement

Xin Chen: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ting Zou:** Writing – review & editing, Visualization, Methodology, Investigation. **Qinglu Zeng:** Writing – review & editing, Methodology, Investigation. **Yubing Chen:** Writing – review & editing, Methodology. **Chengfei Zhang:** Writing – review & editing, Methodology, Investigation. **Shan Jiang:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Guicong Ding:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to sincerely thank the participating children and their guardians for their cooperation in sample collection and invaluable assistance throughout the study.

References

[1] H. Çolak, Ç.T. Dülgergil, M. Dalli, M.M. Hamidi, Early childhood caries update: a review of causes, diagnoses, and treatments, J. Nat. Sci. Biol. Med. 4 (2013) 29–38, https://doi.org/10.4103/0976-9668.107257.
[2] J. Zou, Q. Du, L. Ge, J. Wang, X. Wang, Y. Li, et al., Expert consensus on early childhood caries management, Int. J. Oral Sci. 14 (2022) 35, https://doi.org/10.1038/s41368-022-00186-0.
[3] S.C. Leal, E.M. Takeshita, Early childhood caries, editors, in: S Coelho Leal, EM Takeshita (Eds.), Pediatric restorative dentistry, Springer International Publishing, Cham, 2019, pp. 209–219, https://doi.org/10.1007/978-3-319-93426-6_15.
[4] J.T. Pereira, J.K. Knorst, P.B. Luz, I. Bonfadini, M. Scapinello, F.N. Hugo, et al., Impact of early childhood caries and maternal behaviors on oral health-related quality of life of children, Pesqui. Bras. Odontopediatria Clin. Integr. 20 (2020) 1–13, https://doi.org/10.1590/pboci.2020.065.
[5] D.J. Manton, Child Dental Caries – A Global Problem of Inequality, EClinicalMedicine 1 (2018) 3–4, https://doi.org/10.1016/j.eclinm.2018.06.006.
[6] S.P. Szafrański, J. Slots, M. Stiesch, The human oral phageome, Periodontol. 2000 86 (2021) 79–96, https://doi.org/10.1111/prd.12363.
[7] E. Guerin, C. Hill, Shining Light on Human Gut Bacteriophages, Front. Cell Infect. Microbiol. 10 (2020) 481, https://doi.org/10.3389/fcimb.2020.00481.
[8] V. Lazarevic, K. Whiteson, N. Gaïa, Y. Gizard, D. Hernandez, L. Farinelli, et al., Analysis of the salivary microbiome using culture-independent techniques, J. Clin. Bioinforma 2 (2012) 4, https://doi.org/10.1186/2043-9113-2-4.
[9] D.S.A. Goltsman, C.L. Sun, D.M. Proctor, D.B. DiGiulio, A. Robaczewska, B. C. Thomas, et al., Metagenomic analysis with strain-level resolution reveals fine-scale variation in the human pregnancy microbiome, Genome Res. 28 (2018) 1467–1480, https://doi.org/10.1101/gr.236000.118.
[10] N. Nagata, S. Nishijima, Y. Kojima, Y. Hisada, K. Imbe, T. Miyoshi-Akiyama, et al., Metagenomic Identification of Microbial Signatures Predicting Pancreatic Cancer From a Multinational Study, Gastroenterology 163 (2022) 222–238, https://doi.org/10.1053/j.gastro.2022.03.054.
[11] V.V. Dolja, E.V. Koonin, Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer, Virus. Res. 244 (2018) 36–52, https://doi.org/10.1016/j.virusres.2017.10.020.
[12] M. Norling, O. Karlsson, E. Bongcam-Rudloff, The bioinformatics of viral metagenomics, EMBnetJournal 19 (2013) 68, https://doi.org/10.14806/ej.19.A.661.
[13] S. Roux, D. Páez-Espino, I.M.A. Chen, K. Palaniappan, A. Ratner, K. Chu, et al., IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses, Nucleic Acids Res. 49 (2021) D764–D775, https://doi.org/10.1093/nar/gkaa946.
[14] P. Tang, C. Chiu, Metagenomics for the discovery of novel human viruses, Future Microbiol. 5 (2010) 177–189, https://doi.org/10.2217/fmb.09.120.
[15] B.T. Tierney, Y. Tan, A.D. Kostic, C.J. Patel, Gene-level metagenomic architectures across diseases yield high-resolution microbiome diagnostic indicators, Nat. Commun. 12 (2021) 2907, https://doi.org/10.1038/s41467-021-23029-8.
[16] J. Wang, H. Jia, Metagenome-wide association studies: fine-mining the microbiome, Nat. Rev. Microbiol. 14 (2016) 508–522, https://doi.org/10.1038/nrmicro.2016.83.
[17] A. Edlund, T.M. Santiago-Rodriguez, T.K. Boehm, D.T. Pride, Bacteriophage and their potential roles in the human oral cavity, J. Oral Microbiol. 7 (2015) 27423, https://doi.org/10.3402/jom.v7.27423.
[18] S.P. Szafrański, M. Kilian, I. Yang, G. Bei der Wieden, A. Winkel, J. Hegermann, et al., Diversity patterns of bacteriophages infecting Aggregatibacter and Haemophilus species across clades and niches, ISMe J. 13 (2019) 2500–2522, https://doi.org/10.1038/s41396-019-0450-8.
[19] M.J. de la Cruz Peña, F. Martinez-Hernandez, I. Garcia-Heredia, M. Lluesma Gomez, Ò. Fornas, M Martinez-Garcia, Deciphering the human virome with single-virus genomics and metagenomics, Viruses. 10 (2018), https://doi.org/10.3390/v10030113.
[20] J. Wang, Y. Gao, F. Zhao, Phage-bacteria interaction network in human oral microbiome: human oral virome, Environ. Microbiol. 18 (2016) 2143–2158, https://doi.org/10.1111/1462-2920.12923.
[21] N.N. Al-Hebshi, D. Baraniya, T. Chen, J. Hill, S. Puri, M. Tellez, et al., Metagenome sequencing-based strain-level and functional characterization of supragingival microbiome associated with dental caries in children, J. Oral Microbiol. 11 (2019) 1557986, https://doi.org/10.1080/20002297.2018.1557986.
[22] Amado, F., Calheiros-Lobo, M.J., Ferreira, R., Vitorino, R. (2019). Sample treatment for saliva proteomics. In: Capelo-Martínez, JL. (eds) Emerging Sample Treatments in Proteomics. Advances in Experimental Medicine and Biology, vol 1073. Springer, Cham. doi:10.1007/978-3-030-12298-0_2.
[23] T. Zuo, Y. Sun, Y. Wan, Y.K. Yeoh, F. Zhang, C.P. Cheung, et al., Human-Gut-DNA Virome variations across geography, ethnicity, and urbanization, Cell Host. Microbe 28 (2020) 741–751.e4, https://doi.org/10.1016/j.chom.2020.08.005.
[24] A. Gautam, Phenol-chloroform DNA isolation method, editor, in: A Gautam (Ed.), DNA and RNA isolation techniques for non-experts, Springer International Publishing, Cham, 2022, pp. 33–39, https://doi.org/10.1007/978-3-030-94230-4_3.
[25] A. Modi, S. Vai, D. Caramelli, M. Lari, The Illumina sequencing protocol and the NovaSeq 6000 system, Methods Mol. Biol. 2242 (2021) 15–42, https://doi.org/10.1007/978-1-0716-1099-2_2.
[26] S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics. 34 (2018) i884–i890, https://doi.org/10.1093/bioinformatics/bty560.
[27] M. Kh, N. Y, J. M, A. N, W. Hf, K Wj, Centromere reference models for human chromosomes X and Y satellite arrays, Genome Res. 24 (2014), https://doi.org/10.1101/gr.159624.113.
[28] L. B, S Sl, Fast gapped-read alignment with Bowtie 2, Nat. Methods 9 (2012), https://doi.org/10.1038/nmeth.1923.
[29] D. Li, C.M. Liu, R. Luo, K. Sadakane, T.W. Lam, MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph, Bioinformatics. 31 (2015) 1674–1676, https://doi.org/10.1093/bioinformatics/btv033.
[30] J. Guo, B. Bolduc, A.A. Zayed, A. Varsani, G. Dominguez-Huerta, T.O. Delmont, et al., VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses, Microbiome 9 (2021) 37, https://doi.org/10.1186/s40168-020-00990-y.
[31] J. Ren, K. Song, C. Deng, N.A. Ahlgren, J.A. Fuhrman, Y. Li, et al., Identifying viruses from metagenomic data using deep learning, Quant. Biol. 8 (2020) 64–77, https://doi.org/10.1007/s40484-019-0187-4.

[32] J. Ren, N.A. Ahlgren, Y.Y. Lu, J.A. Fuhrman, F. Sun, VirFinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data, Microbiome 5 (2017) 69, https://doi.org/10.1186/s40168-017-0283-5.
[33] F.A.B. Von Meijenfeldt, K. Arkhipova, D.D. Cambuy, F.H. Coutinho, B.E. Dutilh, Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT, Genome Biol. 20 (2019) 217, https://doi.org/10.1186/s13059-019-1817-x.
[34] D. Hyatt, G.L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer, L.J. Hauser, Prodigal: prokaryotic gene recognition and translation initiation site identification, BMC. Bioinformat. 11 (2010) 119, https://doi.org/10.1186/1471-2105-11-119.
[35] S. Roux, E.M. Adriaenssens, B.E. Dutilh, E.V. Koonin, A.M. Kropinski, M. Krupovic, et al., Minimum information about an uncultivated virus genome (MIUViG), Nat. Biotechnol. 37 (2019) 29–37, https://doi.org/10.1038/nbt.4306.
[36] B. B, R. K, D Hg, Sensitive protein alignments at tree-of-life scale using DIAMOND, Nat. Methods 18 (2021), https://doi.org/10.1038/s41592-021-01101-x.
[37] E.W. Sayers, J. Beck, E.E. Bolton, D. Bourexis, J.R. Brister, K. Canese, et al., Database resources of the national center for biotechnology information, Nucleic. Acids. Res. 49 (2020) D10–D17, https://doi.org/10.1093/nar/gkaa892.
[38] E. Drula, M.L. Garron, S. Dogan, V. Lombard, B. Henrissat, N. Terrapon, The carbohydrate-active enzyme database: functions and literature, Nucleic. Acids. Res. 50 (2022) D571–D577, https://doi.org/10.1093/nar/gkab1045.
[39] J. Huerta-Cepas, D. Szklarczyk, D. Heller, A. Hernández-Plaza, S.K. Forslund, H. Cook, et al., eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses, Nucleic. Acids. Res. 47 (2019) D309–D314, https://doi.org/10.1093/nar/gky1085.
[40] H. Ogata, S. Goto, K. Sato, W. Fujibuchi, H. Bono, M. Kanehisa, KEGG: kyoto Encyclopedia of genes and genomes, Nucleic. Acids. Res. 27 (1999) 29–34, https://doi.org/10.1093/nar/27.1.29.
[41] J.C. Pons, D. Paez-Espino, G. Riera, N. Ivanova, N.C. Kyrpides, M. Llabrés, VPF-Class: taxonomic assignment and host prediction of uncultivated viruses based on viral protein families, Bioinformatics. 37 (2021) 1805–1813, https://doi.org/10.1093/bioinformatics/btab026.
[42] S.H. Katsanis, N. Katsanis, Molecular genetic testing and the future of clinical genomics, Nat. Rev. Genet. 14 (2013) 415–426, https://doi.org/10.1038/nrg3493.
[43] L. Liu, Y. Li, S. Li, N. Hu, Y. He, R. Pong, et al., Comparison of next-generation sequencing systems, J. Biomed. Biotechnol. 2012 (2012) 251364, https://doi.org/10.1155/2012/251364.
[44] A. Betts, D.R. Gifford, R.C. MacLean, K.C. King, Parasite diversity drives rapid host dynamics and evolution of resistance in a bacteria-phage system: growth and resistance with parasite diversity, Evolution. 70 (2016) 969–978, https://doi.org/10.1111/evo.12909.
[45] Y. Liu, S.G. Daniel, H.E. Kim, H. Koo, J. Korostoff, F. Teles, et al., Addition of cariogenic pathogens to complex oral microflora drives significant changes in biofilm compositions and functionalities, Microbiome 11 (2023), https://doi.org/10.1186/s40168-023-01561-7.
[46] D.T. Pride, J. Salzman, M. Haynes, F. Rohwer, C. Davis-Long, R.A. White, et al., Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome, ISMe J. 6 (2012) 915–926, https://doi.org/10.1038/ismej.2011.169.
[47] V. Pérez-Brocal, A. Moya, The analysis of the oral DNA virome reveals which viruses are widespread and rare among healthy young adults in Valencia (Spain), PLoS. One 13 (2018) e0191867, https://doi.org/10.1371/journal.pone.0191867.
[48] A.S. Hemadi, R. Huang, Y. Zhou, J. Zou, Salivary proteins and microbiota as biomarkers for early childhood caries risk assessment, Int. J. Oral Sci. 9 (2017) e1, https://doi.org/10.1038/ijos.2017.35.
[49] D. Willner, M. Furlan, R. Schmieder, J.A. Grasis, D.T. Pride, D.A. Relman, et al., Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity, in: Proceedings of the National Academy of Sciences of the United States of America 108, 2011, pp. 4547–4553, https://doi.org/10.1073/pnas.1000089107.
[50] I. Struzycka, The oral microbiome in dental caries, Pol. J. Microbiol. 63 (2014) 127–135.
[51] A.L. Delisle, C.A. Rostkowski, Lytic bacteriophages ofStreptococcus mutans, Curr. Microbiol. 27 (1993) 163–167, https://doi.org/10.1007/BF01576015.
[52] M. Dalmasso, Haas E de, H. Neve, R. Strain, F.J. Cousin, S.R. Stockdale, et al., Isolation of a novel phage with activity against streptococcus mutans biofilms, PLoS. One 10 (2015) e0138651, https://doi.org/10.1371/journal.pone.0138651.
[53] H. Ben-Zaken, R. Kraitman, S. Coppenhagen-Glazer, L. Khalifa, S. Alkalay-Oren, D. Gelman, et al., Isolation and characterization of streptococcus mutans phage as a possible treatment agent for caries, Viruses. 13 (2021) 825, https://doi.org/10.3390/v13050825.
[54] S.T. Abedon, Size does matter – distinguishing bacteriophages by genome length (and 'breadth'), Microbiol. Aust. 32 (2011) 95, https://doi.org/10.1071/MA11095.
[55] R.E. Lenski, B.R. Levin, Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities, Am. Nat. 125 (1985) 585–602, https://doi.org/10.1086/284364.
[56] M. León, R. Bastías, Virulence reduction in bacteriophage resistant bacteria, Front. Microbiol. 06 (2015) 343, https://doi.org/10.3389/FMICB.2015.00343. –343.
[57] P. Baker, P.J. Hill, B.D. Snarr, N. Alnabelseya, M.J. Pestrak, M.J. Lee, et al., Exopolysaccharide biosynthetic glycoside hydrolases can be utilized to disrupt and prevent Pseudomonas aeruginosa biofilms, Sci. Adv. 2 (2016) e1501632, https://doi.org/10.1126/sciadv.1501632.
[58] C.A. Arbour, R. Nagar, H.M. Bernstein, S. Ghosh, Y. Al-Sammarraie, H. C. Dorfmueller, et al., Defining early steps in Bacillus subtilis biofilm biosynthesis, mBio 14 (2023) e0094823, https://doi.org/10.1128/mbio.00948-23.
[59] D. Fleming, L. Chahin, K. Rumbaugh, Glycoside hydrolases degrade polymicrobial bacterial biofilms in wounds, Antimicrob. Agents Chemother 61 (2017) e01998, https://doi.org/10.1128/AAC.01998-16. -16.
[60] R. Parameswaran, M.P. Soumya, K.M. Nampoothiri, Direct utilization and conversion of raw starch to exopolysaccharides by a newly isolated amylolytic Streptococcus sp, J. Biotechnol. 371–372 (2023) 22–32, https://doi.org/10.1016/j.jbiotec.2023.06.002.
[61] A. Acharya, Y. Chan, S. Kheur, L.J. Jin, R.M. Watt, N. Mattheos, Salivary microbiome in non-oral disease: a summary of evidence and commentary, Arch. Oral Biol. 83 (2017) 169–173, https://doi.org/10.1016/j.archoralbio.2017.07.019.
[62] S.R. Abeles, R. Robles-Sikisaka, M. Ly, A.G. Lum, J. Salzman, T.K. Boehm, et al., Human oral viruses are personal, persistent and gender-consistent, ISMe J. 8 (2014) 1753, https://doi.org/10.1038/ismej.2014.31.
[63] E. Lee, S. Park, S. Um, S. Kim, J. Lee, J. Jang, et al., Microbiome of saliva and plaque in children according to age and dental caries experience, Diagnostics. 11 (2021), https://doi.org/10.3390/diagnostics11081324.
[64] K. Eriksson, G. Fei, A. Lundmark, D. Benchimol, L. Lee, Y.O.O. Hu, et al., Periodontal health and oral microbiota in patients with rheumatoid arthritis, JCM 8 (2019) 630, https://doi.org/10.3390/jcm8050630.
[65] C.W. Ku, D. Shivani, J.Q.T. Kwan, S.L. Loy, C. Erwin, K.K.K. Ko, et al., Validation of self-collected buccal swab and saliva as a diagnostic tool for COVID-19, Internat. J. Infect. Dis. 104 (2021) 255–261, https://doi.org/10.1016/j.ijid.2020.12.080.
[66] X. Fan, B.A. Peters, D. Min, J. Ahn, R.B. Hayes, Comparison of the oral microbiome in mouthwash and whole saliva samples, PLoS. One 13 (2018) e0194729, https://doi.org/10.1371/journal.pone.0194729.