cancers

MDPI

Review

Oral Microbiome in Oral Cancer Research from Sampling to Analysis: Strategies, Challenges, and Recommendations

Kelly Yi Ping Liu [1,2,3], Andrew Huang [1], Catherine Pepin [2], Ya Shen [1], Phoebe Tsang [1] and Catherine F. Poh [1,2,*]

[1] Department of Oral Biological and Medical Sciences, Faculty of Dentistry, University of British Columbia, 2199 Wesbrook Mall, Vancouver, BC V6T 1Z3, Canada; ahhuang@student.ubc.ca (A.H.); yashen@dentistry.ubc.ca (Y.S.); phoebe.tsang.is@gmail.com (P.T.)
[2] Department of Basic and Translational Research, BC Cancer Research Institute, 675 West 10th Avenue, Vancouver, BC V5Z 1L3, Canada; catherinepepin08@gmail.com
[3] School of Biomedical Engineering, University of British Columbia, 6088 University Boulevard, Vancouver, BC V6T 1Z3, Canada
* Correspondence: cpoh@dentistry.ubc.ca; Tel.: +1-604-675-8000

Simple Summary

The human mouth contains many types of bacteria, the balance of which can shift, leading to "dysbiosis", which may contribute to the development of oral cancer. Emerging studies are beginning to uncover the role of the oral microbiome and microbial changes in oral cancer risk and progression. However, current research methods vary widely, making it difficult to compare results across studies or identify reliable biomarkers for clinical applications. This review evaluates the different ways oral microbiome samples are obtained, including saliva, mouth rinse, swabs, and tissue, and how these choices affect findings. We highlight how DNA extraction and sequencing methods can introduce bias. We also emphasize the need for standardization and validation protocols, along with transparent reporting, to reduce methodological variability and enable reproducible, clinically interpretable oral microbiome data.

Abstract

The oral microbiome has become an emerging focus of oral cancer research, with growing evidence linking microbial communities to disease development, progression, and prognosis. However, there is limited consensus on optimal sampling strategies, storage methods, and analytical approaches. This narrative review critically evaluates current strategies for sampling, preservation, DNA extraction, sequencing, and data analysis in oral microbiome research related to oral cancer. We compared commonly used sampling methods, including saliva, oral rinse, swab, brush, and tissue biopsy, and reviewed preservation conditions, extraction kits, sequencing platforms, and analytical pipelines reported in recent oral microbiome studies. Sampling approaches affect microbial yield and site specificity. Saliva and oral rinse samples are convenient and noninvasive but may dilute lesion-specific microbial signals, whereas lesion-directed swabbing or brushing yields greater microbial biomass and biological relevance. Preservation media and storage temperature significantly influence microbial stability, and DNA extraction methods vary in their ability to remove host DNA. Although 16S rRNA gene sequencing remains the most common approach, shotgun metagenomics offers higher resolution and function insights but is still limited by clinical applicability. Differences in data pre- and post-processing models and normalization strategies further contribute to inconsistent microbial profiles. Given that oral mucosal sites differ markedly in structure and microenvironment, careful consideration is required to ensure that collected samples accurately represent the biological question being addressed.

Check for updates

Academic Editor: Yasusei Kudo

Received: 12 December 2025
Revised: 28 December 2025
Accepted: 28 December 2025
Published: 31 December 2025

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Methodological consistency across all workflow stages—from collection to analysis—is essential to generate reproducible, high-quality data and to enable reliable translation of oral microbiome research into clinical applications for cancer detection and risk assessment. Together, these insights provide a framework to guide future study design and support the development of clinically applicable microbiome-based biomarkers.

Keywords: oral microbiome; oral cancer; dysbiosis; methodology; sampling; storage; sequencing; analysis pipeline; clinical applicability

1. Introduction

The human oral cavity is home to one of the most diverse and dynamic microbial ecosystems in the body, consisting of bacteria, fungi, archaea, and protozoa. Collectively known as the oral microbiota, these communities play essential roles in maintaining oral and systemic health through metabolic, immunomodulatory, and protective functions [1,2]. Disruption of this ecosystem, known as dysbiosis, has been implicated not only in oral diseases such as dental caries, periodontal disease, and oral cancer [3,4], but also in a wide range of systemic conditions, including cardiovascular disease, diabetes, respiratory infections, rheumatoid arthritis, adverse pregnancy outcomes, and even neurodegenerative disorders [5–9].

Although many risk factors for oral cancer, such as tobacco and alcohol, are well established, a substantial proportion of cases arise in individuals without these exposures, suggesting additional etiologic factors. The oral microbiome has emerged as one such potential contributor, with studies linking specific microbial signatures to chronic inflammation, immune modulation, epithelial barrier disruption, and carcinogenic metabolites [10,11]. Shifts in microbial communities may precede or accompany malignant transformation, creating opportunities for early detection and risk stratification. Characterizing the oral microbiome in oral cancer could therefore provide new insights into disease mechanisms and support the development of microbiome-based diagnostic and prognostic tools.

Beyond taxonomic profiling, emerging multi-omic and functional approaches are increasingly used to interrogate microbial function and biological activity. These technologies enable researchers to examine not only which organisms are present, but also their metabolic pathways and interactions with epithelial and immune responses. Despite growing interest in the oral microbiome as a biomarker for oral cancer, there remains a limited consensus on standardized methodologies for sample collection, processing, and analysis. Such inconsistencies have hindered progress and validation of oral microbiome research.

The purpose of this narrative review is to provide a comprehensive methodological evaluation that can guide the design of future oral microbiome research in premalignant oral diseases and oral cancer. We examine how variations in biosample types, collection methods, preservation conditions, DNA extraction procedures, sequencing strategies, and data processing workflows influence microbial profiles. Our aim is to identify the major sources of technical variability that may limit reproducibility across studies and hinder the discovery of robust clinically applicable biomarkers.

Review Design/Literature Search Strategy

This narrative review gathers recent evidence on sampling, preservation, DNA extraction, sequencing strategies, and data analysis in oral microbiome research related to oral premalignant lesions and oral cancer. We conducted targeted literature searches in PubMed and Web of Science for English-language articles published up to 30 August 2025 using combination of the following keywords: “oral microbiome”, “oral cancer”, “oral squamous

cell carcinoma", "oral potentially malignant lesions", "sampling", "saliva", "rinse", "swab", "brush", "tissue", "DNA extraction", "16S rRNA", "shotgun metagenomics", "pipeline", and "database". Additional relevant studies were identified by screening reference lists of key articles and recent reviews. Studies were included if they were relevant to our review objectives, with a focus on the methodology. Only human studies were included. The titles and abstracts of all articles were reviewed, and those relevant to our goals were selected for a full-text review. After reviewing the full texts, articles that were relevant were included in various sections.

2. Sample Types and Collection Techniques

Differences in sample types and collection techniques can lead to variability in microbial biomass, taxonomic composition, and downstream analysis. Therefore, the type of samples should depend on the research goal. In hypothesis-driven studies, where the goal is to test specific microbial-associated diseases, the target community can often be reasonably anticipated, and sampling can be tailored accordingly. In contrast, explorative studies that aim to broadly characterize community composition should prioritize representativeness and reproducibility to minimize technical bias. In oral cancer research, most microbiome-oriented studies remain exploratory, seeking to uncover potential microbial biomarkers. Consequently, variability in sampling techniques has become a major methodological consideration. Understanding these differences is essential for interpreting existing findings and for guiding standardization efforts. Table 1 provides an overview of current studies investigating the oral microbiome in the context of oral malignant and premalignant oral lesions with different sampling methodologies [12,13].

Table 1. Sample types and sampling techniques in studies on oral precancers and cancer.

Study (Year)	Sampling Method(s)	Sample Size	Study Goal	Key Findings
Sasaki et al., 2005 [14]	Tissue & dental plaque	46	Compare plaque and tissue in OSCC patients vs. controls	Tissue and plaque in OSCC show distinct microbial community shifts
Pushalkar et al., 2012 [15]	Tissue	20	Profile tumour-associated microbiome in OSCC	OSCC tissues enriched in anaerobes like *Fusobacterium nucleatum*
Yang et al., 2018 [16]	Oral rinse	248	Analyze oral rinse differences between OSCC and controls	Significant changes in community composition in OSCC group
Yang et al., 2021 [17]	Saliva (unstimulated) & tissue	42	Characterize OSCC-associated changes in saliva and tissue samples	Tumour tissues enriched with anaerobes including *Fusobacterium nucleatum*
Gopinath et al., 2021 [18]	Saliva, swabbing, & biopsy tissue	94	Compare microbiomes across niches in oral cancer patients	Surface vs. deep tissue niche differences, and saliva-tissue comparisons in terms of composition and metabolic pathways
Shitozawa et al., 2022 [19]	Rinse & swabbing	10	Compare intra-individual differences of microbiota between lesion and normal and swabbing, vs. oral rinse	Lesion-specific dysbiosis observed, highlighting importance of site-specific sampling

Table 1. *Cont.*

Study (Year)	Sampling Method(s)	Sample Size	Study Goal	Key Findings
Wright et al., 2023 [20]	Brushing & swabbing	90	Compare lesional vs. normal site microbiome in OSCC cases	Lesional sites showed enrichment of *Fusobacterium* and reduced diversity vs. contralateral controls
Mäkinen et al., 2023 [21]	Saliva (stimulated)	236	Survey salivary microbiome in OSCC vs. controls	Salivary profiles showed distinct community shifts in OSCC, with increased pathogenic taxa
Deo et al., 2025 [22]	Saliva	50	Compare normal, OPMD, OSCC microbiomes	Gram-negative and periodontal taxa increased in OSCC; unique taxa found in OSCC group
Intini et al., 2025 [23]	Saliva	66	Differentiate OL, PVL, OSCC microbiomes	OL shows higher richness, OSCC lowest, beta diversity distinguishes OL from PVL/OSCC

Abbreviation(s): OSCC, oral squamous cell carcinoma; OPMDs, oral potentially malignant disorders; OL, oral leukoplakia; PVL, proliferative verrucous leukoplakia.

2.1. Saliva

Saliva is the most commonly collected oral sample in microbiome studies due to its ease of collection, noninvasiveness, and relatively high DNA yield. Since oral surfaces are bathed in saliva, these samples capture microbes from mucosal surfaces as well as free-floating bacteria, providing a broad overview of the oral microbiota. Unstimulated saliva is collected naturally through drooling, spitting, or cotton absorption [12]. This approach avoids stimulation-induced changes but may yield low volumes in patients with hyposalivation, such as those undergoing or recovering from oral cancer treatments, which can potentially alter microbial composition and diversity. In contrast, stimulated saliva, obtained by chewing paraffin or gum, tasting sour substances, or even stimulating imagery [12], provides larger sample volumes in a shorter time. While several studies have reported comparable results between stimulated and unstimulated saliva [24,25], others have documented compositional differences [26].

Saliva collection requires minimal technical expertise; therefore, it is suitable for self-administration and large-scale population studies. Moreover, the salivary microbiome has been shown to be relatively stable over short periods—from hours to weeks—and even up to a year [27,28]. Intra-individual profiles of both diversity and community composition remain more similar within a person than between individuals, with inter-individual differences dominating the variance [28]. Short-term, high-frequency sampling has revealed eating-associated fluctuations in specific taxa [29]. In contrast, in hospitalized or medically ill patients exposed to intensive chemotherapy and prolonged antibiotic use, repeated saliva sampling shows pronounced temporal instability. In addition, greater intra-patient variability is associated with adverse outcomes [30]. Together, these data suggest that saliva is a suitable method for temporal microbiome profiling within individuals, provided that sampling is standardized with respect to time of day, diet, and oral hygiene, and should account for medical conditions. In oral cancer research, saliva has been used to detect general microbial shifts or inflammatory signatures, but may be less informative for detecting early, localized changes in the mucosal microbiome, which are more effectively captured through site-directed sampling [17,19,31].

2.2. Oral Rinse

Oral rinse, similar to saliva, is an easy and noninvasive approach for collecting a broad, non-site-specific overview of the oral microbiome. The technique typically involves participants rinsing and/or gargling a defined volume of sterile collection medium or mouthwash for a fixed duration before spitting into a sterile tube. The rinsing process mechanically dislodges microorganisms from mucosal surfaces. In most studies, participants are instructed to alternate between swishing and gargling in five-second intervals for a total of 30 s before spitting the sample into a sterile container [32,33]. An alternate protocol involves swishing without gargling, such as the one described by Yu et al. [34]. Swishing alone may provide a more targeted profile of the oral cavity rather than the oropharyngeal microbiome [35]. While the choice of collection medium may influence taxonomic composition, mouthwash-based media have been shown to yield reliable oral microbiome profiles [33,35].

2.3. Exfoliation by Swabbing, Brushing, and Scrapping

Swabbing, brushing, and scraping enable the collection of site-specific microbial communities by targeting microbes attached to specific epithelial surfaces rather than free-floating in saliva. Swabbing is minimally invasive and typically performed with cotton or nylon-flocked swabs. Brushing, which uses cytobrushes or sponge brushes, is a more aggressive technique that dislodges biofilms and exfoliates cells. While both methods are effective for collecting niche-representative samples, brushing is generally preferred for lesion sampling as it yields higher microbial biomass and can recover microorganisms embedded deeper within biofilms [19]. In a study investigating the progression of oral epithelial dysplasia, researchers used cytology brushes (Innovatek Inc., Delta, BC, Canada) to exfoliate cells directly from oral lesions and obtained informative microbiome results [20]. Comparable success with brushing techniques has been demonstrated in other mucosal sites, such as the cervix and lung, where they provide high-quality microbial DNA for sequencing [36–39]. However, brushing may introduce higher proportions of host epithelial cells and can cause discomfort in sensitive tissues.

Scraping with a metal spatula is a more invasive method but provides enough material for microscopic analysis of epithelial-associated microorganisms, especially when samples are spread directly onto glass slides. This method enables fluorescent in situ hybridization (FISH) with DNA or RNA probes that target specific 16S rRNA sequences, allowing direct detection of microorganisms adhering to the epithelial cells [40,41]. When combined with sequencing data, this method can reveal the presence, relative abundance, and proximity of specific taxa to host cells, providing complementary insights into host–microbe interactions at the mucosal surface [40].

2.4. Biopsy Tissue Samples

Tissue samples are particularly valuable in oral cancer research as they capture microorganisms in direct association with precancerous or cancerous epithelium, potentially revealing site- or depth-specific microbial patterns not detectable with other sampling methods. Beyond enabling spatial localization of microbes within the tissue architecture, biopsy-based microbiome studies have identified unique bacteria associated with chronic inflammation and different clinical stages of cancer [15,17]. Similar to scraping samples, biopsies can be analyzed using FISH with probes specific to microbial antigens [42,43], allowing the visualization of microorganisms within the tissue microenvironment. This approach facilitates investigation of host–microbe interactions, localization of specific taxa, and validation of sequencing results. However, the use of biopsy samples in microbiome research is limited by clinical availability, ethical considerations, and sample size limitations.

In summary, the choice of sampling method influences the biological signal captured because different oral niches harbour distinct microbial communities. Saliva and oral rinses primarily reflect the pooled, planktonic microbiome, representing bacteria shed from multiple mucosal and dental surfaces. These non-site-specific samples are well-suited for population-level screening, epidemiologic studies, and assessing global microbial shifts, but they may dilute or mask localized dysbiosis associated with early lesions. In contrast, swabs and brushing samples adhere to biofilm directly, providing a more localized view of the microbial community. Lesion-directed brushing or swabbing enhances the detection of biomarkers associated with dysplasia, inflammation, or tumour-associated microbiota, making them more appropriate for risk stratification, early detection, and tumour microenvironment studies. Tissue biopsies offer the most spatially resolved information, enabling analysis of bacteria that physically associate with epithelial layers, invasion, or co-localize with inflammatory infiltrates. Thus, tissue samples are most suitable for studying host–microbe interactions, depth-specific ecology, and tumour-associated communities.

3. Sample Preservation and Storage Condition

Collected samples are usually stabilized using preservation media to prevent bacterial outgrowth, DNA degradation, and loss of microbial integrity. Effective preservation should minimize changes in microbial composition between collection and downstream processing. A variety of preservation media have been used for oral microbiome samples (Table 2). However, relatively few studies have directly compared their effects on microbial stability and sequencing outcomes. Recognizing and controlling for these influences is essential for producing reliable and comparable data [44,45].

Table 2. Oral microbiome studies comparing preservation buffer among different sampling techniques.

Study (Year)	Collection Method	Preservation Buffer/Kit	Sample Size	Key Findings
Luo et al., (2016) [32]	Plaque, saliva, swabbing	OMNIgene ORAL, Liquid Dental Transfer Medium (LDTM)	10	OMNIgene yielded higher diversity than LDTM; OMNIgene stable up to 1 week at room temperature, LDTM stable across multiple temperatures but with different composition
Vogtmann et al., (2019) [33]	Rinse & saliva (Scope® mouthwash)	Scope® mouthwash, OMNIgene ORAL	100	Mouthwash samples remained stable at room temperature for 4 days. Microbial community profiles differed between methods.
Zhou et al., (2019) [46]	Dental plaque	75% ethanol vs. bead solution	20	Storage medium influenced microbial recovery more than extraction method; ethanol and bead solutions maintained community structure well
Yano et al., (2020) [35]	Rinse (saline)	OMNIgene ORAL, Scope® mouthwash, non-ethanol rinse, Saccomano's fixative	40	OMNIgene yielded distinct, stable profiles while Saccomano's samples showed the greatest divergence in diversity and taxonomic composition
Bang et al., (2023) [47]	Saliva, rinse (saline)	DNA/RNA Shield, OMNIgene ORAL saliva kit	3	Oral rinses showed higher diversity; OMNIgene, DNA/RNA Shield and simple collections produced similar profiles; taxonomic differences were not significant
Yano et al., (2023) [48]	Oral rinse	Ethanol-free, ethanol-based mouthwash	40	Both stable up to 10 days; minor taxa shifts by formulation; similar composition overall

3.1. Commercial Stabilization Kits

Commercial kits, such as OMNIgene ORAL (DNA Genotek Inc., Ottawa, ON, Canada) and DNA/RNA Shield (Zymo Research, Tustin, CA, USA), contain proprietary buffers designed to lyse host cells and preserve microbial DNA at room temperature. These kits are available in multiple formats, including saliva and site-specific sampling devices, and are widely used in oral microbiome research. While no universal "gold standard" exists for oral sample preservation, OMNIgene saliva kits are frequently used as a benchmark in comparative studies. According to the manufacturer's data, DNA stability can be maintained up to 21 days, whereas swab-based kits can stabilize samples up to 30 days at room temperature [49]. The long stability period, combined with ease of use, makes commercial kits attractive for large cohort studies, multi-site collaboration, or remote and self-collection scenarios. However, comparative data between commercial kits remain limited.

3.2. Fixative-Type Preservation Media

Fixatives such as Saccomano's fixative are designed to preserve cellular morphology but may alter microbial composition. The fixative, containing 50% ethanol and 2% Carbowax, was originally developed for sputum cytology and later adapted for microbiome studies [50]. While effective at maintaining cellular morphology, comparative studies show that samples preserved in Saccomanno's fixative produce microbial profiles that differ from those obtained with other preservation methods, exhibiting increased beta-divergence, reduced alpha diversity, and altered taxonomic composition compared to commercial kits or ethanol-based oral rinses [35]. These differences are likely due to the relatively low ethanol concentration and the absence of dedicated microbial stabilizers, which permit bacterial overgrowth during storage, compromising the accuracy of microbiome profiling [35].

Another transport medium, the liquid dental transfer medium (LDTM), contains buffered mineral salts, sodium thioglycolate, and cysteine to maintain microbial viability while minimizing overgrowth during transport [51]. While OMNIgene samples typically show higher within-sample diversity [32], LDTM has demonstrated stability across a range of storage temperatures, making it a practical option for studies involving shipping or delayed processing. However, comparative studies show that LDTM samples harbor significantly different community compositions compared with OMNIgene samples, highlighting concerns for buffer-specific bias.

3.3. Ethanol- and Non-Ethanol-Based Mouth Washes

Compared with commercial kits and fixatives, ethanol and non-ethanol mouthwashes offer convenient and low-cost alternatives for sample storage. However, their chemical properties can influence the stability and accuracy of downstream microbial profiles. While generally stable at room temperature, studies have reported that samples stored in ethanol-based mouthwash begin to exhibit shifts in microbial composition after extended storage. For example, one study found that after four days at room temperature, *Firmicutes* increased while *Bacteroidetes*, *Proteobacteria*, and *Fusobacteria* decreased, relative to the sample preserved in OMNIgene [33]. In another study, 95% ethanol has been shown to preserve community profiles at room temperature for up to 8 weeks [52]. Although higher ethanol concentrations may improve stability, these findings have not been systematically validated for oral microbiome samples and should therefore be interpreted with caution.

Ethanol- and non-ethanol-based mouthwashes differ in how they interact with microorganisms. Ethanol reduces microbial viability through protein denaturation and membrane permeabilization, but it does not kill or lyse most bacteria because their cell walls typically remain intact. Consequently, ethanol-tolerant taxa may survive and proliferate slowly during storage, introducing post-collection shifts in microbial composition.

In contrast, non-ethanol preservation buffers often contain surfactants and chemical stabilizers, such as SDS in OMNIgene, that rapidly disrupt cell membranes, inactivate microbial metabolic activity, and suppress post-collection overgrowth [35]. These formulations are therefore more effective at maintaining the true microbial profile and may provide more consistent results than ethanol-based solutions.

Regardless of the chosen medium, it is important that the type of transport or preservation solution be consistent across all samples to ensure that the observed microbial differences reflect disease states rather than technical artifacts.

3.4. Storage Conditions

Beyond the choice of collection buffer, sample storage conditions may be another factor influencing sample stability and the integrity of oral microbiome data. Short-term storage at room temperature or 4 °C is often unavoidable, but the degree of microbial change depends largely on the medium used. Current evidence suggests that with appropriate preservatives (e.g., OMNIgene, DNA/RNA Shield), oral samples can remain stable at ambient temperature for 1–2 weeks [53]. Nonetheless, best practice from broader microbiome research recommends rapid freezing at −20 °C, and ideally −80 °C, particularly for studies requiring high sensitivity or long-term storage. Repeated freeze–thaw cycles should be minimized, as they can degrade DNA and reduce recovery of low-abundance organisms [52–55].

4. DNA Extraction and Sequencing

4.1. Microbial DNA Extraction

The goal of microbial DNA extraction is to efficiently recover high-quality nucleic acids while minimizing contamination from host DNA. Oral samples contain a large proportion of human epithelial cells and saliva; therefore, extraction protocols must balance effective depletion of host DNA and effective microbial cell lysis. A wide range of extraction methods and commercial kits are available, each differing in how they disrupt the cells and isolate the DNA. For example, the QIAamp DNA Microbiome kit (Qiagen) combines mechanical disruption, such as bead-beating, with host DNA depletion and enrichment of bacterial DNA. In contrast, the PureLink Microbiome DNA purification kit (Thermo Fisher Scientific, Waltham, MA, USA) recovers both microbial and host DNA, though its performance and suitability for integrated host–microbiome analyses in oral samples have not yet been evaluated [55].

Comparative studies have shown that extraction protocols incorporating mechanical disruption, such as bead-beating or vortex-assisted lysis, yield higher DNA quantities and capture greater microbial diversity compared to enzymatic or chemical lysis alone, particularly bacteria with robust cell walls [56]. Therefore, careful consideration should also account for DNA extraction methods that may bias relative abundance estimates. New host-depletion techniques that aim to improve microbial biomass have emerged. For example, osmotic lysis combined with propidium monoazide treatment has been used to enrich microbial DNA in both fresh and frozen saliva [57–59].

4.2. Sequencing Strategies

4.2.1. 16S rRNA Gene Sequencing Preparation

16S rRNA amplicon sequencing is one of the most common methods for characterizing the oral microbiome. This approach targets hypervariable regions (V1–V9) of the bacterial 16S rRNA gene via PCR amplification, with primer selection strongly influencing taxonomic resolution. For example, Zheng et al. found that the V1–V3 regions captured higher phylotype richness compared to the V3–V4 regions. On the other hand, Kool et al. found

that V3–V4 and V4–V5 regions yielded more reproducible results across replicates, although their findings were based on a synthetic community [60,61]. While most oral cancer studies focus on V3–V4 or V4–V5 [15,62] regions rather than V1–V3 [20,63,64], there is no consensus on which hypervariable regions offer optimal coverage and resolution, as results also depend on the sequencing platform and downstream analysis (Table 3).

Table 3. Sequencing strategies in oral microbiome studies of precancer and cancer.

Study (Year)	Sequencing Platform; 16S rRNA Region	General Processing Pipeline (Clustering)	Taxonomic Assign-ment	Depth (Reads)	Rarefied	Key Findings
Pushalkar et al., 2012 [15]	Sanger sequencing; V4–V5	Greengenes NAST (OUT)	HOMD	~95	No (rar-efaction only)	Tumour tissues are enriched with anaerobes including *Fusobacterium nucleatum* and *Prevotella* species
Yang et al., 2021 [17]	Illumina MiSeq; V3–V4	DADA2 (ASV)	RDP, HOMD	~23,645	Not reported	Tumour and saliva showed distinct shifts. *Fusobacterium* is enriched in tumour tissue
Gopinath et al., 2021 [18]	Illumina MiSeq; V3–V5	DADA2 (ASV)	HOMD	Not reported	Not reported	Tumour surface enriched in Prophyromonas, *Fusobacteria*, and *Neisseria*. Deeper tissue enriched in *Prevotella* and *Treponema*. Saliva reflected tissue composition but had higher metabolic pathway diversity
Shitozawa et al., 2022 [19]	Sanger sequencing; V3–V5	CD-HIT-EST (OTU)	RDP	~192	No (rar-efaction only)	Lesion microbiota is less diverse and distinct from contralateral sites. Oral rinses captured broader but less site-specific diversity
Mäkinen et al., 2023 [21]	Illumina MiSeq; V4	UNOISE3 (ASV)	HOMD	Not reported	Not reported	Post-treatment saliva showed reduced diversity but persistence of pathogenic taxa associated with OSCC.
Wright et al., 2023 [20]	Illumina MiSeq; V1–V3	Cutadapt, UNOISE3 (ASV)	HOMD	~7907	Yes	OEDs showed distinct microbial profiles with reduced diversity and enrichment of pro-inflammatory taxa

Abbreviation(s): ASV, amplicon sequence variant; OTUs, operational taxonomic units; HOMD, Human Oral Microbiome Database; RDP, the Ribosomal Database Project; OED, oral epithelial dysplasia; OSCC, oral squamous cell carcinoma.

4.2.2. Shotgun Metagenomic Sequencing Preparation

Shotgun metagenomic sequencing provides comprehensive profiling of the genetic repertoire of a sample, enabling both species- and strain-level identification and functional annotation. Ganly et al. used metagenomic shotgun sequencing on oral rinse samples from nonsmoking OSCC patients and demonstrated distinct taxonomic and metabolic signatures compared with nonsmoking controls, including enrichment of pathways involved in vitamin, heme, and carbohydrate metabolism [65]. Among OSCC patients, Liu et al. further characterized microbiome variation across different depths of invasion, revealing progressive shifts in bacterial composition and microbial functions associated with more

invasive tumours. Unlike 16S rRNA amplicon sequencing, shotgun sequencing is not limited to primer selection bias and can simultaneously detect bacteria, fungi, archaea, and viruses. However, shotgun sequencing suffers from substantial host DNA contamination, adapter ligation bias during library preparation, uneven genome coverage, and taxonomic misclassification resulting from incomplete or inaccurate reference databases [59]. Strategies to reduce host contamination include selective lysis of host cells, enzymatic digestion of host DNA, or targeting highly methylated eukaryotic DNA to enrich the remaining microbial DNA [59,66].

Sequencing strategies determine both the taxonomic resolution and functional insights in oral microbiome research. 16S rRNA gene sequencing remains a widely used and cost-effective approach, requiring relatively less DNA (10–50 ng DNA per sample) and modest sequencing depth. It is supported by well-established bioinformatics pipelines with limitations on primer bias, reduced taxonomic resolution, and inability to directly infer microbial function [67]. In contrast, shotgun metagenomic sequencing requires more DNA materials (100–1000 ng DNA per sample) and deeper sequencing coverage but provides species-to-strain-level precision and enables direct functional annotation of genes and pathways [68]. Given that shotgun sequencing is considerably more expensive, many studies use 16S rRNA sequencing for the full cohort and perform shotgun sequencing only on a smaller subset of samples. This targeted approach allows researchers to validate the 16S-derived taxonomic patterns and obtain additional functional and strain-level insights that 16S cannot provide [69].

4.3. Sequencing Platform Selection

The choice of sequencing platform can also influence oral microbiome data, as technologies differ in read length, throughput, and error rate. Awareness of these differences is critical when comparing results across studies or integrating multiple datasets.

Currently, Illumina platforms remain the most widely used platform in oral microbiome research as they provide high throughput with low error rates at relatively low cost. The MiSeq system is commonly used for 16S rRNA amplicon sequencing, whereas the NovaSeq platform supports deeper shotgun metagenomic sequencing for large-scale studies [70,71]. Compared with older platforms, such as pyrosequencing, Illumina's higher accuracy, greater throughput, and mature bioinformatics pipelines make it well-suited for taxonomic profiling and diversity studies [72].

Third-generation sequencing technologies, such as PacBio HiFi and Oxford Nanopore, generate long-read sequencing that enables full-length 16S rRNA gene sequencing and improved species- or strain-level resolution in human oral microbiome samples [73]. PacBio HiFi, in particular, has improved species-level assignment compared to short-read sequencing [73]. However, PacBio systems generally offer lower throughput and higher per-sample costs, limiting their feasibility for large cohorts. Studies have reported mixed results regarding diversity estimation, with some studies showing no significant differences compared with Illumina data, while others report altered alpha diversity and relative abundance profiles [74]. While Oxford Nanopore sequencing provides real-time, portable sequencing capabilities, the higher error rates can reduce taxonomic accuracy unless robust error-correction algorithms are applied.

Comparative studies suggest that while estimates of alpha diversity and taxonomic abundance can vary among platforms, overall biological conclusions are consistent across methods. Studies must therefore balance cost, accuracy, read length, and data compatibility when selecting a sequencing platform [75].

5. Data Processing and Analysis

High-throughput sequencing produces large and complex datasets that require careful processing to ensure that downstream analyses accurately reflect the true microbial community. Key steps typically include quality filtering, chimera removal, denoising or clustering of sequence variants, and taxonomic classification using curated reference databases. Differences in how these steps are implemented can lead to substantial variability in reported microbial diversity and composition. As summarized in Table 3, studies vary widely in their sequence processing and taxonomic assignment methods, underscoring the importance of harmonized data-processing pipelines to minimize analytical bias.

5.1. Quality Control, Denoising, and Clustering

Processing raw sequencing data into biologically interpretable features involves a series of structured computational steps. These steps determine how microbial abundance is represented and are critical for ensuring that downstream analyses accurately reflect the underlying community. Processing typically begins with quality control that removes technical artifacts and low-quality reads. This includes adapter and primer trimming, filtering sequences by quality scores or expected error thresholds, removing sequences of inappropriate length, and identifying and removing chimeric reads [76,77]. Tools such as QIIME2, Mothur, USEARCH, FASTQC, Trimmomatic, and Cutadapt are commonly used for these steps. For shotgun metagenomics, tools such as FastQC [78] are commonly used to generate visual summaries of read quality, while Trimmomatic and Cutadapt are used for adapter trimming and quality filtering based on Phred scores [77,79].

After quality filtering, sequences are processed into either amplicon sequence variants (ASVs) or operational taxonomic units (OTUs), which are the core units to represent microbial features in downstream analysis. ASV-based approaches, such as DADA2 and Deblur, use statistical error models to correct sequencing errors and infer exact biological sequences. These methods detect single-nucleotide differences and provide higher resolution and taxonomic precision. In contrast, OTU-based approaches, such as CD-HIT, USEARCH/VSEARCH, or UNOISE3, cluster sequences into groups based on a predefined similarity threshold (typically $\geq$ 97–99.8%). While widely used in earlier studies and useful for integrating legacy datasets, OTU clustering can identify fine-grained ecological differences between closely related taxa [80,81]. Following clustering, sequences are collapsed into a feature table that summarizes the abundance of each ASV and OTU across samples.

For shotgun metagenomic sequencing, an additional preprocessing step is required that removes human reads. Oral samples often contain a large proportion of host DNA, which can obscure microbial signals. Host depletion is performed by aligning reads against the human reference genome and retaining only unmapped reads. Tools such as KneadData, BWA [82], or newer host-contamination removal frameworks such as HoCoRT are widely used for this purpose [83,84].

5.2. Reference Database for Taxonomic Assignment

Taxonomic assignment is typically performed using reference databases. For 16S rRNA gene sequencing, curated databases align reads to known taxa. The Human Oral Microbiome Database (HOMD) [85] is frequently used in oral microbiome studies because it provides oral-specific taxonomy and curated reference sequences from cultured and uncultured species. Broader databases such as SILVA [81], RDP [80], or Greengenes2 [86] offer wider phylogenetic coverage across diverse environments but may lack depth or accurate annotation for some oral-specific taxa. Some studies leverage multiple databases for validation or to achieve better genus- or species-level resolution. Reporting the version

of the reference database and assignment parameters is therefore crucial for reproducibility and cross-study comparability.

For shotgun metagenomics sequencing, taxonomic and functional annotation require a more complex workflow because of the greater volume and diversity of data. After initial quality control, reads are often assembled into contiguous sequences (contigs) using assemblers, such as MEGAHIT [87] or metaSPAdes [88], that are optimized for large metagenomic datasets. Benchmarking studies have shown that no single assembler is universally optimal. Performance varies depending on factors such as sequencing depth, sample complexity, and computational resources [89,90]. Modular frameworks such as MetAMOS [91] provide flexibility to evaluate multiple assembly and annotation pipelines, optimizing performance for a given dataset [92]. In some cases, analysis of unassembled reads is preferred, particularly for highly diverse or low-coverage samples where assembly quality may be limited. For the taxonomic classification of such reads, k-mer-based algorithms such as Kraken and its derivatives [93–95] have demonstrated high accuracy and computational efficiency.

5.3. Normalization and Data Transformation

Interpretation of sequencing data requires careful normalization to account for variation in sequencing depth and library size across samples. No single method is universally optimal, and each introduces unique assumptions and biases. Recent benchmarking studies suggest that using a consensus-based approach by applying multiple methods and focusing on reproducible signals strengthens the robustness of microbial profiles [96]. The simplest normalization converts raw counts into relative abundances, expressing each taxon as a percentage of total reads per sample. Although this approach remains widely used, it is limited by the compositional nature of microbiome data, where all taxa within a sample must sum to 100% [97]. Consequently, an apparent increase in the relative abundance of one taxa reflect a decrease in another, even if its absolute abundance has not changed. Such interdependence between taxa can create spurious correlations and distort downstream analyses [98]. To address this, compositional-aware methods are increasingly recommended because they account for these dependences and provide more reliable inferences [96].

5.4. Sampling Depth and Filtering

Sequencing depth varies considerably between samples and platforms, making normalization and filtering essential to ensure fair comparisons. Rarefaction is a common approach for standardizing sequencing depth across samples, where samples are subsampled to the lowest acceptable read depth within a study. This procedure reduces biases introduced by uneven sequencing, as samples with more reads capture more rare taxa [99]. Rarefaction curves can guide this process, with plateaus indicating that sequencing depth is sufficient to capture most microbial diversity. However, rarefaction also discards valid reads, reduces statistical power, and may increase false positives or negatives. Alternatively, model-based normalization methods, such as the negative binomial models in DESeq2, have been developed [100] for sparse microbial count data using the median-of-ratios method. More recent compositional-aware tools, including ALDEx2, ANCOM-BC, corncob, and songbird, further account for zero-inflation and compositional constraints, offering improved control of false discoveries in microbiome datasets [101–103]. These approaches adjust for both library size and biological variability without discarding data, thereby improving reproducibility and sensitivity, particularly useful when sample sizes are limited or when subtle microbial shifts are being investigated.

Filtering low-abundance or low-relevance features is another critical-step to reduce noise [104]. Such thresholds are often determined empirically, but a prevalence threshold of 10% (i.e., retaining taxa detected in at least 10% of samples) is commonly applied to

ensure more robust and reproducible results [96]. Relative abundance thresholds, typically between 0.01% and 0.1% per sample, help to remove extremely rare taxa that could distort diversity or differential abundance analyses [104]. However, overly stringent thresholds risk discarding rare but biologically meaningful taxa.

6. Discussion

The oral microbiome is a dynamic and diverse ecosystem influenced by host factors, environmental exposures, and disease processes. Understanding and standardizing sampling strategies is critical for the growing number of studies investigating oral premalignant and OSCC. Saliva and oral rinses offer convenient, noninvasive options suitable for population-level screenings and for investigating systemic associations. However, these pooled samples may dilute site-specific microbial signals, limiting their ability to detect localized dysbiosis associated with early lesion development. In contrast, site-directed sampling methods such as exfoliation and biopsies are the most appropriate when the research question targets lesion-associated microbial shifts, though they are invasive and require greater technical expertise. Collectively, there is no single sampling approach that is universally superior; rather, the optimal method must align with the biological question in study design.

In addition, maintaining consistency in preservation media, DNA extraction methods, and sequencing workflows is essential to minimize technical biases. Standardized protocols and transparent reporting will be critical to improving reproducibility, enabling cross-study comparison, and translating oral microbiome profiles into clinically meaningful applications. Recent reviews have further emphasized how methodological choices can shape biological interpretation. For example, Zhi et al. used shotgun metagenomics on tissue-based samples and identified *Streptococcus*, rather than *Fusobacterium*, as a dominant genus associated with oral cancer, contrasting with findings from other saliva-based 16S rRNA studies [105]. Such discrepancies highlight how differences in sampling sites, sequencing platforms, and analytical pipelines can bias which taxa emerge as the significant biomarkers for OSCC, emphasizing the need for methodological standardization under the same research goals.

Another important aspect is that the oral cavity comprises multiple anatomically and functionally distinct niches [106]. Each site, influenced by factors such as oxygen availability, moisture, nutrient gradients, pH, and host immune interactions, supports a unique microbial community. Understanding the clinical relevance of these niches is essential in oral cancer research, where lesion-associated dysbiosis is often site-specific. For example, the rough, papillary surface of the tongue harbors a community enriched in anaerobes, whereas the gingival crevice, affected frequently by periodontal disease, supports a markedly different microbial profile [107]. Paired sampling of a lesion and its contralateral normal mucosa provides a powerful within-subject design that controls for inter-individual variability and helps to highlight microbial shifts characterized in the abnormal lesion site.

The oral cavity is also easily influenced by short-term behavioral and environmental factors; pre-collection protocols are therefore essential for minimizing variability [29,108,109]. Standard recommendations include avoiding eating, drinking, smoking, chewing gum, and performing oral hygiene for at least 60 min prior to sampling. Time of day should also be standardized when possible, as circadian and eating-associated microbial fluctuation can affect community composition. Clear instructions and consistent training of participants and clinicians help to reduce pre-analytical variability. Recommendations for contamination control include the use of sterile collection instruments, gloves, and single-use consumables. Consistent adherence to pre-collection and contamination control procedures helps ensure

that observed microbial differences reflect true biological variation rather than extraneous technical artifacts. In addition, it is important to recognize that sampling approaches, particularly invasive biopsy, may influence a patient's willingness to participate, thereby potentially limiting recruitment and reducing cohort size. Even though longitudinal and repeated sampling can capture microbial dynamics associated with disease progression, accounting for temporal variation remains challenging.

Quality control is fundamental for reliable oral microbiome research. Positive controls, such as synthetic microbial communities, environmental samples from participants, or in vitro microbial models, can verify the integrity of extraction and sequencing workflows [110]. Negative controls, such as reagent blanks or extraction controls, are equally important, especially in low-biomass oral samples [110,111]. Given that sequencing is highly sensitive and can amplify even small amounts of contaminant DNA, negative controls are critical for detecting and addressing contamination. Studies should therefore transparently report their DNA extraction methods and contamination control measures to ensure reproducibility and comparability across studies [110]. Host DNA contamination, particularly in oral lesion or brush samples, remains a significant challenge [112]. In addition to the current methods involving mechanical disruption and selective lysis of mammalian cells, emerging approaches aim to further increase microbial signal through microbial enrichment and hybridization kits to selectively bind and remove host DNA before library construction. However, no host-depletion strategy is entirely unbiased. Overly aggressive depletion can inadvertently remove microbial taxa with fragile cell walls or low biomass.

While most oral microbiome studies rely on DNA-based sequencing to infer community composition and potential function, emerging RNA-based and other functional omics approaches are beginning to reveal active microbial pathways and host–microbe interactions. These methods, which capture transcriptionally active or metabolically expressed signals, can provide more dynamic insights into the biological role of the microbiome. However, they remain technically challenging. The review by Hu et al. highlights the promise of advanced assay technologies (e.g., mass spectrometry, multiplex platforms, nanobiosensors) and AI-driven analytics for saliva-based biomarker detection, as well as the persistent challenges of standardization and biological variability [113]. Incorporating multi-omic approaches, such as integrating metagenomic, transcriptomic, metabolomic, and host genomic data, can deepen biological interpretation and help identify mechanistic links between microbial pathways and host responses.

Clinical translation requires a robust framework to connect microbial alterations, reliable metrics to identify reliable biomarkers, and interventional studies to test whether modifying the microbiome can alter outcomes. Thus, a rigorous study design is essential to move from observational associations to clinical applications.

7. Conclusions

The oral microbiome represents a promising frontier in understanding the pathogenesis and progression of oral premalignant lesions and OSCC. Emerging evidence suggests that microbial signatures may serve as biomarkers for early detection, risk stratification, and monitoring treatment response. However, the translational potential of these findings remains constrained. Our review highlights key areas of inconsistency across studies, ranging from sampling and preservation methods to DNA extraction, sequencing strategies, and bioinformatics processing, that contribute to divergent microbial profiles and hinder reproducibility. Addressing these challenges through standardized and transparent methodologies will be critical to generate reliable and comparable data. Ultimately, oral microbiome profiling holds the potential to deliver noninvasive, clinically actionable

biomarkers for early cancer detection and surveillance, as well as to identify therapeutic targets for microbiome modulation aimed at improving patient outcomes.

Author Contributions: K.Y.P.L., A.H. and C.F.P. conceptually designed the structure of the review. K.Y.P.L., A.H., C.P., P.T. and C.F.P. performed the literature review. K.Y.P.L., A.H. and C.P. drafted the manuscript. K.Y.P.L., A.H., Y.S., P.T. and C.F.P. discussed and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Kelly YP Liu was supported by the Canadian Institutes of Health Research Fellowship; Andrew Huang and Catherine Pepin were supported by the BC Cancer Foundation (Oral Cancer Better Outcomes).

Institutional Review Board Statement: No applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This review did not generate any new data.

Acknowledgments: We thank Steven Hallam and Avery Nanoon for their constructive guidance and valuable feedback.

Conflicts of Interest: The authors have no conflicts of interest to declare.

Abbreviations

The following abbreviations are used in this manuscript:

OSCC	Oral squamous cell carcinoma
OPMD	Oral potentially malignant disorders
OL	Oral leukoplakia
PVL	Proliferative verrucous leukoplakia
FISH	Fluorescence in situ hybridization
LDTM	Liquid dental transfer medium
ASV	Amplicon sequence variant
OTU	Operational taxonomic units
HOMD	Human Oral Microbiome Database
RDP	Ribosomal Database Project
OED	Oral epithelial dysplasia

References

1. Kilian, M.; Chapple, I.L.; Hannig, M.; Marsh, P.D.; Meuric, V.; Pedersen, A.M.; Tonetti, M.S.; Wade, W.G.; Zaura, E. The oral microbiome—An update for oral healthcare professionals. *Br. Dent. J.* **2016**, *221*, 657–666.
2. Zaura, E.; Nicu, E.A.; Krom, B.P.; Keijser, B.J. Acquiring and maintaining a normal oral microbiome: Current perspective. *Front. Cell. Infect. Microbiol.* **2014**, *4*, 85. [CrossRef]
3. Di Stefano, M.; Polizzi, A.; Santonocito, S.; Romano, A.; Lombardi, T.; Isola, G. Impact of Oral Microbiome in Periodontal Health and Periodontitis: A Critical Review on Prevention and Treatment. *Int. J. Mol. Sci.* **2022**, *23*, 5142. [CrossRef]
4. Zhang, J.S.; Chu, C.-H.; Yu, O.Y. Oral Microbiome and Dental Caries Development. *Dent. J.* **2022**, *10*, 184. [CrossRef]
5. Lalla, E.; Papapanou, P.N. Diabetes mellitus and periodontitis: A tale of two common interrelated diseases. *Nat. Rev. Endocrinol.* **2011**, *7*, 738–748. [CrossRef]
6. Han, Y.W.; Wang, X. Mobile microbiome: Oral bacteria in extra-oral infections and inflammation. *J. Dent. Res.* **2013**, *92*, 485–491.
7. Tonetti, M.S.; Van Dyke, T.E. working group 1 of the joint EFPAAPw. Periodontitis and atherosclerotic cardiovascular disease: Consensus report of the Joint EFP/AAP Workshop on Periodontitis and Systemic Diseases. *J. Periodontol.* **2013**, *84*, S24–S29.
8. Mikuls, T.R.; Payne, J.B.; Yu, F.; Thiele, G.M.; Reynolds, R.J.; Cannon, G.W.; Markt, J.; McGowan, D.; Kerr, G.S.; Redman, R.S.; et al. Periodontitis and Porphyromonas gingivalis in patients with rheumatoid arthritis. *Arthritis Rheumatol.* **2014**, *66*, 1090–1100. [CrossRef]
9. Scannapieco, F.A. Role of oral bacteria in respiratory infection. *J. Periodontol.* **1999**, *70*, 793–802. [CrossRef]
10. Eun, Y.G.; Lee, J.W.; Kim, S.W.; Hyun, D.W.; Bae, J.W.; Lee, Y.C. Oral microbiome associated with lymph node metastasis in oral squamous cell carcinoma. *Sci. Rep.* **2021**, *11*, 23176. [CrossRef]
11. Irfan, M.; Delgado, R.Z.R.; Frias-Lopez, J. The Oral Microbiome and Cancer. *Front. Immunol.* **2020**, *11*, 591088. [CrossRef]

12. Lu, H.; Zou, P.; Zhang, Y.; Zhang, Q.; Chen, Z.; Chen, F. The sampling strategy of oral microbiome. *iMeta* **2022**, *1*, e23. [CrossRef]
13. Špiljak, B.; Ozretić, P.; Andabak Rogulj, A.; Lončar Brzak, B.; Brailo, V.; Škerlj, M.; Vidović Juras, D. Oral Microbiome Research in Biopsy Samples of Oral Potentially Malignant Disorders and Oral Squamous Cell Carcinoma and Its Challenges. *Appl. Sci.* **2024**, *14*, 11405. [CrossRef]
14. Sasaki, M.; Yamaura, C.; Ohara-Nemoto, Y.; Tajika, S.; Kodama, Y.; Ohya, T.; Harada, R.; Kimura, S. *Streptococcus anginosus* infection in oral cancer and its infection route. *Oral Dis.* **2005**, *11*, 151–156. [CrossRef]
15. Pushalkar, S.; Ji, X.; Li, Y.; Estilo, C.; Yegnanarayana, R.; Singh, B.; Li, X.; Saxena, D. Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. *BMC Microbiol.* **2012**, *12*, 144. [CrossRef]
16. Yang, C.-Y.; Yeh, Y.-M.; Yu, H.-Y.; Chin, C.-Y.; Hsu, C.-W.; Liu, H.; Huang, P.-J.; Hu, S.-N.; Liao, C.-T.; Chang, K.-P.; et al. Oral Microbiota Community Dynamics Associated With Oral Squamous Cell Carcinoma Staging. *Front. Microbiol.* **2018**, *9*, 862. [CrossRef]
17. Yang, K.; Wang, Y.; Zhang, S.; Zhang, D.; Hu, L.; Zhao, T.; Zheng, H. Oral Microbiota Analysis of Tissue Pairs and Saliva Samples From Patients With Oral Squamous Cell Carcinoma—A Pilot Study. *Front. Microbiol.* **2021**, *12*, 719601. [CrossRef]
18. Gopinath, D.; Menon, R.K.; Wie, C.C.; Banerjee, M.; Panda, S.; Mandal, D.; Behera, P.K.; Roychoudhury, S.; Kheur, S.; Botelho, M.G.; et al. Differences in the bacteriome of swab, saliva, and tissue biopsies in oral cancer. *Sci. Rep.* **2021**, *11*, 1181. [CrossRef]
19. Shitozawa, Y.; Haro, K.; Ogawa, M.; Miyawaki, A.; Saito, M.; Fukuda, K. Differences in the microbiota of oral rinse, lesion, and normal site samples from patients with mucosal abnormalities on the tongue. *Sci. Rep.* **2022**, *12*, 16839. [CrossRef]
20. Wright, R.J.; Pewarchuk, M.E.; Marshall, E.A.; Murray, B.; Rosin, M.P.; Laronde, D.M.; Zhang, L.; Lam, W.L.; Langille, M.G.I.; Rock, L.D. Exploring the microbiome of oral epithelial dysplasia as a predictor of malignant progression. *BMC Oral Health* **2023**, *23*, 206. [CrossRef]
21. Mäkinen, A.I.; Pappalardo, V.Y.; Buijs, M.J.; Brandt, B.W.; Mäkitie, A.A.; Meurman, J.H.; Zaura, E. Salivary microbiome profiles of oral cancer patients analyzed before and after treatment. *Microbiome* **2023**, *11*, 171. [CrossRef]
22. Deo, P.N.; Deshmukh, R.S.; Gaike, A.H.; Christopher, A.; Gujare, M.; Inamdar, M. Oral microbiome profiles in oral potentially malignant disorders and oral cancer—A diagnostic perspective. *J. Oral Maxillofac. Pathol.* **2025**, *29*, 87–97. [CrossRef]
23. Intini, R.; Balsells, S.; Bagan, L.; Fortuna, G.; Sroussi, H.; Bagan, J. Comparative analysis of oral microbiome in saliva samples of oral leukoplakia, proliferative leukoplakia and oral squamous cell carcinoma. *Front. Oral Health* **2025**, *6*, 1600090. [CrossRef]
24. Belstrøm, D.; Holmstrup, P.; Bardow, A.; Kokaras, A.; Fiehn, N.-E.; Paster, B.J. Comparative analysis of bacterial profiles in unstimulated and stimulated saliva samples. *J. Oral Microbiol.* **2016**, *8*, 30112. [CrossRef]
25. Jo, R.; Nishimoto, Y.; Umezawa, K.; Yama, K.; Aita, Y.; Ichiba, Y.; Murakami, S.; Kakizawa, Y.; Kumagai, T.; Yamada, T.; et al. Comparison of oral microbiome profiles in stimulated and unstimulated saliva, tongue, and mouth-rinsed water. *Sci. Rep.* **2019**, *9*, 16124. [CrossRef]
26. Gomar-Vercher, S.; Simón-Soro, A.; Montiel-Company, J.M.; Almerich-Silla, J.M.; Mira, A. Stimulated and unstimulated saliva samples have significantly different bacterial profiles. *PLoS ONE* **2018**, *13*, e0198021. [CrossRef]
27. Belstrøm, D.; Holmstrup, P.; Bardow, A.; Kokaras, A.; Fiehn, N.-E.; Paster, B.J. Temporal Stability of the Salivary Microbiota in Oral Health. *PLoS ONE* **2016**, *11*, e0147472. [CrossRef]
28. Cameron, S.J.S.; Huws, S.A.; Hegarty, M.J.; Smith, D.P.M.; Mur, L.A.J. The human salivary microbiome exhibits temporal stability in bacterial diversity. *FEMS Microbiol. Ecol.* **2015**, *91*, fiv091. [CrossRef]
29. Hu, Y.; Amir, A.; Huang, X.; Li, Y.; Huang, S.; Wolfe, E.; Weiss, S.; Knight, R.; Xu, Z.Z. Diurnal and eating-associated microbial patterns revealed via high-frequency saliva sampling. *Genome Res.* **2022**, *32*, 1112–1123. [CrossRef]
30. Galloway-Pena, J.R.; Smith, D.P.; Sahasrabhojane, P.; Wadsworth, W.D.; Fellman, B.M.; Ajami, N.J.; Shpall, E.J.; Daver, N.; Guindani, M.; Petrosino, J.F.; et al. Characterization of oral and gut microbiome temporal variability in hospitalized cancer patients. *Genome Med.* **2017**, *9*, 21. [CrossRef]
31. Belstrom, D.; Holmstrup, P.; Fiehn, N.E.; Kirkby, N.; Kokaras, A.; Paster, B.J.; Bardow, A. Salivary microbiota in individuals with different levels of caries experience. *J. Oral Microbiol.* **2017**, *9*, 1270614. [CrossRef]
32. Luo, T.; Srinivasan, U.; Ramadugu, K.; Shedden, K.A.; Neiswanger, K.; Trumble, E.; Li, J.J.; McNeil, D.W.; Crout, R.J.; Weyant, R.J.; et al. Effects of Specimen Collection Methodologies and Storage Conditions on the Short-Term Stability of Oral Microbiome Taxonomy. *Appl. Environ. Microbiol.* **2016**, *82*, 5519–5529. [CrossRef]
33. Vogtmann, E.; Chen, J.; Kibriya, M.G.; Amir, A.; Shi, J.; Chen, Y.; Islam, T.; Eunes, M.; Ahmed, A.; Naher, J.; et al. Comparison of Oral Collection Methods for Studies of Microbiota. *Cancer Epidemiol. Biomark. Prev.* **2019**, *28*, 137–143. [CrossRef]
34. Yu, G.; Phillips, S.; Gail, M.H.; Goedert, J.J.; Humphrys, M.; Ravel, J.; Ren, Y.; Caporaso, N.E. Evaluation of Buccal Cell Samples for Studies of Oral Microbiota. *Cancer Epidemiol. Biomark. Prev.* **2017**, *26*, 249–253. [CrossRef]
35. Yano, Y.; Hua, X.; Wan, Y.; Suman, S.; Zhu, B.; Dagnall, C.L.; Hutchinson, A.; Jones, K.; Hicks, B.D.; Shi, J.; et al. Comparison of Oral Microbiota Collected Using Multiple Methods and Recommendations for New Epidemiologic Studies. *mSystems* **2020**, *5*, e00156-20. [CrossRef]

36. Martinez-Medina, J.N.; Ghazisaeedi, F.; Kramer, C.; Ziegler, J.F.; McParland, V.; Mönch, P.W.; Siegmund, B.; Jarquín-Díaz, V.H.; Fulde, M.; Forslund-Startceva, S.K. Mucosal washes are useful for sampling intestinal mucus-associated microbiota despite low biomass. *Gut Microbes* **2025**, *17*, 2464296. [CrossRef]
37. Noonan, A.J.C.; Myers, R.; Bartolomeu, C.; Borden, S.A.; Lam, S.; Hallam, S.J. Correlation Mapping of the Human Respiratory Tract Microbiome Reveals Lung-Specific Indicators and Subnetworks. In Proceedings of the American Thoracic Society 2024 International Conference, San Diego, CA, USA, 17–22 May 2024; American Thoracic Society: New York, NY, USA, 2024.
38. Tsay, J.-C.J.; Wu, B.G.; Badri, M.H.; Clemente, J.C.; Shen, N.; Meyn, P.; Li, Y.; Yie, T.-A.; Lhakhang, T.; Olsen, E.; et al. Airway Microbiota Is Associated with Upregulation of the PI3K Pathway in Lung Cancer. *Am. J. Respir. Crit. Care Med.* **2018**, *198*, 1188–1198. [CrossRef]
39. Yi, X.; Gao, J.; Wang, Z. The human lung microbiome—A hidden link between microbes and human health and diseases. *iMeta* **2022**, *1*, e33. [CrossRef]
40. Mark Welch, J.L.; Rossetti, B.J.; Rieken, C.W.; Dewhirst, F.E.; Borisy, G.G. Biogeography of a human oral microbiome at the micron scale. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E791–E800. [CrossRef]
41. Wilbert, S.A.; Mark Welch, J.L.; Borisy, G.G. Spatial Ecology of the Human Tongue Dorsum Microbiome. *Cell Rep.* **2020**, *30*, 4003–4015.e4003. [CrossRef]
42. Gu, J.; Wang, H.; Zhang, M.; Xiong, Y.; Yang, L.; Ren, B.; Huang, R. Application of Fluorescence In Situ Hybridization (FISH) in Oral Microbial Detection. *Pathogens* **2022**, *11*, 1450. [CrossRef]
43. Nejman, D.; Livyatan, I.; Fuks, G.; Gavert, N.; Zwang, Y.; Geller, L.T.; Rotter-Maskowitz, A.; Weiser, R.; Mallel, G.; Gigi, E.; et al. The human tumor microbiome is composed of tumor type–specific intracellular bacteria. *Science* **2020**, *368*, 973–980. [CrossRef]
44. Choo, J.M.; Leong, L.E.; Rogers, G.B. Sample storage conditions significantly influence faecal microbiome profiles. *Sci. Rep.* **2015**, *5*, 16350. [CrossRef]
45. Superdock, D.K.; Zhang, W.; Poole, A.C. Processing and storage methods affect oral and gut microbiome composition. *Front. Microbiol.* **2023**, *14*, 1253570. [CrossRef]
46. Zhou, X.; Nanayakkara, S.; Gao, J.L.; Nguyen, K.A.; Adler, C.J. Storage media and not extraction method has the biggest impact on recovery of bacteria from the oral microbiome. *Sci. Rep.* **2019**, *9*, 14968. [CrossRef]
47. Bang, E.; Oh, S.; Ju, U.; Chang, H.E.; Hong, J.-S.; Baek, H.-J.; Kim, K.-S.; Lee, H.-J.; Park, K.U. Factors influencing oral microbiome analysis: From saliva sampling methods to next-generation sequencing platforms. *Sci. Rep.* **2023**, *13*, 10086. [CrossRef]
48. Yano, Y.; Vogtmann, E.; Shreves, A.H.; Weinstein, S.J.; Black, A.; Diaz-Mayoral, N.; Wan, Y.; Zhou, W.; Hua, X.; Dagnall, C.L.; et al. Evaluation of alcohol-free mouthwash for studies of the oral microbiome. *PLoS ONE* **2023**, *18*, e0284956. [CrossRef]
49. DNA Genotek Inc. *OMNIgene Oral—Optimized Solutions for Sample Collection, Nucleic Acid Stabilization and Microbial Profile Snapshot*; Report No.: PD-BR-00205 Issue 4; DNA Genotek Inc.: Ottawa, ON, Canada, 2020.
50. Biognost Ltd. *Saccomanno Fixative—Instructions for Use V2*; Report No.: SF-X, V2-EN2; Biognost: Zagreb, Croatia, 2019.
51. Inc AS. *Liquid Dental Transport—LDT*; Anaerobe Systems: Morgan Hill, CA, USA, 2017.
52. Song, S.J.; Amir, A.; Metcalf, J.L.; Amato, K.R.; Xu, Z.Z.; Humphrey, G.; Knight, R. Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. *mSystems* **2016**, *1*, e00021-16. [CrossRef]
53. Kim, H.J.; Park, D.H.; Han, S.H.; Kim, S.-Y. Optimal storage time and temperature of human oral samples to minimize microbiome changes: A scoping review. *Jpn. Dent. Sci. Rev.* **2024**, *60*, 220–231. [CrossRef]
54. Jenkins, S.V.; Vang, K.B.; Gies, A.; Griffin, R.J.; Jun, S.-R.; Nookaew, I.; Dings, R.P.M. Sample storage conditions induce post-collection biases in microbiome profiles. *BMC Microbiol.* **2018**, *18*, 227. [CrossRef]
55. Pollock, J.; Glendinning, L.; Wisedchanwet, T.; Watson, M. The Madness of Microbiome: Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. *Appl. Environ. Microbiol.* **2018**, *84*, e02627-17. [CrossRef]
56. Teng, F.; Darveekaran Nair, S.S.; Zhu, P.; Li, S.; Huang, S.; Li, X.; Xu, J.; Yang, F. Impact of DNA extraction method and targeted 16S-rRNA hypervariable region on oral microbiota profiling. *Sci. Rep.* **2018**, *8*, 16321.
57. Horz, H.-P.; Scheer, S.; Huenger, F.; Vianna, M.E.; Conrads, G. Selective isolation of bacterial DNA from human clinical specimens. *J. Microbiol. Methods* **2008**, *72*, 98–102. [CrossRef]
58. Liu, G.; Weston, C.Q.; Pham, L.K.; Waltz, S.; Barnes, H.; King, P.; Sphar, D.; Yamamoto, R.T.; Forsyth, R.A. Epigenetic Segregation of Microbial Genomes from Complex Samples Using Restriction Endonucleases HpaII and McrB. *PLoS ONE* **2016**, *11*, e0146064.
59. Marotz, C.A.; Sanders, J.G.; Zuniga, C.; Zaramela, L.S.; Knight, R.; Zengler, K. Improving saliva shotgun metagenomics by chemical host DNA depletion. *Microbiome* **2018**, *6*, 42. [CrossRef] [PubMed]
60. Zheng, W.; Tsompana, M.; Ruscitto, A.; Sharma, A.; Genco, R.; Sun, Y.; Buck, M.J. An accurate and efficient experimental approach for characterization of the complex oral microbiota. *Microbiome* **2015**, *3*, 48. [CrossRef]
61. Kool, J.; Tymchenko, L.; Shetty, S.A.; Fuentes, S. Reducing bias in microbiome research: Comparing methods from sample collection to sequencing. *Front. Microbiol.* **2023**, *14*, 1094800. [CrossRef]
62. Yu, X.; Shi, Y.; Yuan, R.; Chen, Z.; Dong, Q.; Han, L.; Wang, L.; Zhou, J. Microbial dysbiosis in oral squamous cell carcinoma: A systematic review and meta-analysis. *Heliyon* **2023**, *9*, e13198. [CrossRef]

63. Al-Hebshi, N.N.; Nasher, A.T.; Idris, A.M.; Chen, T. Robust species taxonomy assignment algorithm for 16S rRNA NGS reads: Application to oral carcinoma samples. *J. Oral Microbiol.* **2015**, *7*, 28934. [CrossRef]
64. Perera, M.; Al-Hebshi, N.N.; Perera, I.; Ipe, D.; Ulett, G.C.; Speicher, D.J.; Chen, T.; Johnson, N.W. Inflammatory Bacteriome and Oral Squamous Cell Carcinoma. *J. Dent. Res.* **2018**, *97*, 725–732. [CrossRef]
65. Ganly, I.; Hao, Y.; Rosenthal, M.; Wang, H.; Migliacci, J.; Huang, B.; Katabi, N.; Brown, S.; Tang, Y.W.; Pei, Z.; et al. Oral Microbiome in Nonsmoker Patients with Oral Cavity Squamous Cell Carcinoma, Defined by Metagenomic Shotgun Sequencing. *Cancers* **2022**, *14*, 6096. [CrossRef] [PubMed]
66. Feehery, G.R.; Yigit, E.; Oyola, S.O.; Langhorst, B.W.; Schmidt, V.T.; Stewart, F.J.; Dimalanta, E.T.; Amaral-Zettler, L.A.; Davis, T.; Quail, M.A.; et al. AMethod for Selectively Enriching Microbial DNAfrom Contaminating Vertebrate Host DNA. *PLoS ONE* **2013**, *8*, e76096. [CrossRef]
67. Wensel, C.R.; Pluznick, J.L.; Salzberg, S.L.; Sears, C.L. Next-generation sequencing: Insights to advance clinical investigations of the microbiome. *J. Clin. Investig.* **2022**, *132*, e154944. [CrossRef] [PubMed]
68. Bars-Cortina, D.; Ramon, E.; Rius-Sansalvador, B.; Guinó, E.; Garcia-Serrano, A.; Mach, N.; Khannous-Lleiffe, O.; Saus, E.; Gabaldón, T.; Ibáñez-Sanz, G.; et al. Comparison between 16S rRNA and shotgun sequencing in colorectal cancer, advanced colorectal lesions, and healthy human gut microbiota. *BMC Genom.* **2024**, *25*, 730. [CrossRef]
69. Torralba, M.G.; Aleti, G.; Li, W.; Moncera, K.J.; Lin, Y.H.; Yu, Y.; Masternak, M.M.; Golusinski, W.; Golusinski, P.; Lamperska, K.; et al. Oral Microbial Species and Virulence Factors Associated with Oral Squamous Cell Carcinoma. *Microb. Ecol.* **2021**, *82*, 1030–1046. [CrossRef] [PubMed]
70. Abdill, R.J.; Graham, S.P.; Rubinetti, V.; Ahmadian, M.; Hicks, P.; Chetty, A.; McDonald, D.; Ferretti, P.; Gibbons, E.; Rossi, M.; et al. Integration of 168,000 samples reveals global patterns of the human gut microbiome. *Cell* **2025**, *188*, 1100–1118.e1117. [CrossRef]
71. Bharti, R.; Grimm, D.G. Current challenges and best-practice protocols for microbiome analysis. *Brief. Bioinform.* **2021**, *22*, 178–193. [CrossRef]
72. Allali, I.; Arnold, J.W.; Roach, J.; Cadenas, M.B.; Butz, N.; Hassan, H.M.; Koci, M.; Ballou, A.; Mendoza, M.; Ali, R.; et al. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol.* **2017**, *17*, 194. [CrossRef]
73. Buetas, E.; Jordán-López, M.; López-Roldán, A.; D'Auria, G.; Martínez-Priego, L.; De Marco, G.; Carda-Diéguez, M.; Mira, A. Full-length 16S rRNA gene sequencing by PacBio improves taxonomic resolution in human microbiome samples. *BMC Genom.* **2024**, *25*, 310. [CrossRef]
74. Katiraei, S.; Anvar, Y.; Hoving, L.; Berbée, J.F.P.; Van Harmelen, V.; Willems Van Dijk, K. Evaluation of Full-Length Versus V4-Region 16S rRNA Sequencing for Phylogenetic Analysis of Mouse Intestinal Microbiota After a Dietary Intervention. *Curr. Microbiol.* **2022**, *79*, 276. [CrossRef]
75. Biada, I.; Santacreu, M.A.; González-Recio, O.; Ibáñez-Escriche, N. Comparative analysis of Illumina, PacBio, and nanopore for 16S rRNA gene sequencing of rabbit's gut microbiota. *Front. Microbiomes* **2025**, *4*, 1587712. [CrossRef]
76. Jeske, J.T.; Gallert, C. Microbiome Analysis via OTU and ASV-Based Pipelines—A Comparative Interpretation of Ecological Data in WWTP Systems. *Bioengineering* **2022**, *9*, 146. [CrossRef] [PubMed]
77. Breitwieser, F.P.; Lu, J.; Salzberg, S.L. A review of methods and databases for metagenomic classification and assembly. *Brief. Bioinform.* **2019**, *20*, 1125–1136. [CrossRef] [PubMed]
78. Braham Bioinformatics. *FastQC A Quality Control Tool for High Throughput Sequence Data*; Braham Bioinformatics: Cambridge, UK, 2010.
79. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **2011**, *17*, 10. [CrossRef]
80. Maidak, B.L.; Olsen, G.J.; Larsen, N.; Overbeek, R.; McCaughey, M.J.; Woese, C.R. The RDP (Ribosomal Database Project). *Nucleic Acids Res.* **1997**, *25*, 109–110. [CrossRef] [PubMed]
81. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2012**, *41*, D590–D596. [CrossRef]
82. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef]
83. Gao, Y.; Luo, H.; Lyu, H.; Yang, H.; Yousuf, S.; Huang, S.; Liu, Y.X. Benchmarking short-read metagenomics tools for removing host contamination. *Gigascience* **2025**, *14*, giaf004. [CrossRef]
84. Rumbavicius, I.; Rounge, T.B.; Rognes, T. HoCoRT: Host contamination removal tool. *BMC Bioinform.* **2023**, *24*, 371. [CrossRef]
85. Chen, T.; Yu, W.-H.; Izard, J.; Baranova, O.V.; Lakshmanan, A.; Dewhirst, F.E. The Human Oral Microbiome Database: A web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* **2010**, *2010*, baq013. [CrossRef]
86. McDonald, D.; Jiang, Y.; Balaban, M.; Cantrell, K.; Zhu, Q.; Gonazalez, A.; Morton, J.T.; Nicolaou, G.; Parks, D.H.; Karst, S.M.; et al. Greengenes2 unifies microbial data in a single reference tree. *Nat. Biotechnol.* **2024**, *42*, 715–718. [CrossRef]
87. Li, D.; Liu, C.-M.; Luo, R.; Sadakane, K.; Lam, T.-W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. *Bioinformatics* **2015**, *31*, 1674–1676. [CrossRef] [PubMed]

88. Nurk, S.; Meleshko, D.; Korobeynikov, A.; Pevzner, P.A. metaSPAdes: A new versatile metagenomic assembler. *Genome Res.* **2017**, *27*, 824–834. [CrossRef]
89. Bradnam, K.R.; Fass, J.N.; Alexandrov, A.; Baranay, P.; Bechner, M.; Birol, I.; Boisvert, S.; Chapman, J.A.; Chapuis, G.; Chikhi, R.; et al. Assemblathon 2: Evaluating de novo methods of genome assembly in three vertebrate species. *Gigascience* **2013**, *2*, 10. [CrossRef]
90. Quince, C.; Walker, A.W.; Simpson, J.T.; Loman, N.J.; Segata, N. Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* **2017**, *35*, 833–844. [CrossRef] [PubMed]
91. Treangen, T.J.; Koren, S.; Sommer, D.D.; Liu, B.; Astrovskaya, I.; Ondov, B.; Darling, A.E.; Phillippy, A.M.; Pop, M. MetAMOS: A modular and open source metagenomic assembly and analysis pipeline. *Genome Biol.* **2013**, *14*, R2. [CrossRef] [PubMed]
92. Roumpeka, D.D.; Wallace, R.J.; Escalettes, F.; Fotheringham, I.; Watson, M. A Review of Bioinformatics Tools for Bio-Prospecting from Metagenomic Sequence Data. *Front. Genet.* **2017**, *8*, 23. [CrossRef]
93. Lu, J.; Breitwieser, F.P.; Thielen, P.; Salzberg, S.L. Bracken: Estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **2017**, *3*, e104. [CrossRef]
94. Breitwieser, F.P.; Baker, D.N.; Salzberg, S.L. KrakenUniq: Confident and fast metagenomics classification using unique k-mer counts. *Genome Biol.* **2018**, *19*, 198. [CrossRef]
95. Wood, D.E.; Lu, J. Langmead B: Improved metagenomic analysis with Kraken 2. *Genome Biol.* **2019**, *20*, 257. [CrossRef]
96. Nearing, J.T.; Douglas, G.M.; Hayes, M.G.; MacDonald, J.; Desai, D.K.; Allward, N.; Jones, C.M.A.; Wright, R.J.; Dhanani, A.S.; Comeau, A.M.; et al. Microbiome differential abundance methods produce different results across 38 datasets. *Nat. Commun.* **2022**, *13*, 342. [CrossRef] [PubMed]
97. Hawinkel, S.; Mattiello, F.; Bijnens, L.; Thas, O. A broken promise: Microbiome differential abundance methods do not control the false discovery rate. *Brief. Bioinform.* **2019**, *20*, 210–221. [CrossRef] [PubMed]
98. Gloor, G.B.; Macklaim, J.M.; Pawlowsky-Glahn, V.; Egozcue, J.J. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* **2017**, *8*, 2224. [CrossRef]
99. Willis, A.D. Rarefaction, Alpha Diversity, and Statistics. *Front. Microbiol.* **2019**, *10*, 2407. [CrossRef] [PubMed]
100. McMurdie, P.J.; Holmes, S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Comput. Biol.* **2014**, *10*, e1003531. [CrossRef]
101. Fernandes, A.D.; Macklaim, J.M.; Linn, T.G.; Reid, G.; Gloor, G.B. ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. *PLoS ONE* **2013**, *8*, e67019. [CrossRef]
102. Lin, H.; Peddada, S.D. Analysis of compositions of microbiomes with bias correction. *Nat. Commun.* **2020**, *11*, 3514. [CrossRef]
103. Calgaro, M.; Romualdi, C.; Waldron, L.; Risso, D.; Vitulo, N. Assessment of statistical methods from single cell, bulk RNA-seq, and metagenomics applied to microbiome data. *Genome Biol.* **2020**, *21*, 191. [CrossRef]
104. Zhou, R.; Ng, S.K.; Sung, J.J.Y.; Goh, W.W.B.; Wong, S.H. Data pre-processing for analyzing microbiome data—A mini review. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 4804–4815. [CrossRef]
105. Zhi, J.; Liang, Y.; Zhao, W.; Qiao, J.; Zheng, Y.; Peng, X.; Li, L.; Wei, X.; Wang, W. Oral microbiome-derived biomarkers for non-invasive diagnosis of head and neck squamous cell carcinoma. *NPJ Biofilms Microbiomes* **2025**, *11*, 74. [CrossRef]
106. Human Microbiome Project, C. Structure, function and diversity of the healthy human microbiome. *Nature* **2012**, *486*, 207–214. [CrossRef]
107. Baker, J.L.; Mark Welch, J.L.; Kauffman, K.M.; McLean, J.S.; He, X. The oral microbiome: Diversity, biogeography and human health. *Nat. Rev. Microbiol.* **2024**, *22*, 89–104. [CrossRef]
108. Caporaso, J.G.; Lauber, C.L.; Costello, E.K.; Berg-Lyons, D.; Gonzalez, A.; Stombaugh, J.; Knights, D.; Gajer, P.; Ravel, J.; Fierer, N.; et al. Moving pictures of the human microbiome. *Genome Biol.* **2011**, *12*, R50. [CrossRef]
109. Flores, G.E.; Caporaso, J.G.; Henley, J.B.; Rideout, J.R.; Domogala, D.; Chase, J.; Leff, J.W.; Vázquez-Baeza, Y.; Gonzalez, A.; Knight, R.; et al. Temporal variability is a personalized feature of the human microbiome. *Genome Biol.* **2014**, *15*, 531. [CrossRef]
110. Greathouse, K.L.; Sinha, R.; Vogtmann, E. DNA extraction for human microbiome studies: The issue of standardization. *Genome Biol.* **2019**, *20*, 212. [CrossRef] [PubMed]
111. Salter, S.J.; Cox, M.J.; Turek, E.M.; Calus, S.T.; Cookson, W.O.; Moffatt, M.F.; Turner, P.; Parkhill, J.; Loman, N.J.; Walker, A.W. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **2014**, *12*, 87. [CrossRef]
112. Hillmann, B.; Al-Ghalith, G.A.; Shields-Cutler, R.R.; Zhu, Q.; Gohl, D.M.; Beckman, K.B.; Knight, R.; Knights, D. Evaluating the Information Content of Shallow Shotgun Metagenomics. *mSystems* **2018**, *3*, e00069-18. [CrossRef] [PubMed]
113. Hu, C.C.; Wang, S.G.; Gao, Z.; Qing, M.F.; Pan, S.; Liu, Y.Y.; Li, F. Emerging salivary biomarkers for early detection of oral squamous cell carcinoma. *World J. Clin. Oncol.* **2025**, *16*, 103803. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.