**Taylor & Francis**
Taylor & Francis Group

ORIGINAL ARTICLE

# Rapid specific detection of oral bacteria using Cas13-based SHERLOCK

Jett Liu[a]*, Camden Carmichael[a]*, Hatice Hasturk[b], Wenyuan Shi[a] and Batbileg Bor [a,c]

[a]Department of Microbiology, The Forsyth Institute, Cambridge, MA, USA; [b]Center for Clinical and Translational Research, The Forsyth Institute, Cambridge, MA, USA; [c]Department of Oral Medicine, Infection and Immunity, Harvard School of Dental Medicine, Boston, MA, USA

**ABSTRACT**

Decades of ongoing research has established that oral microbial communities play a role in oral diseases such as periodontitis and caries. Yet the detection of oral bacteria and the profiling of oral polymicrobial communities currently rely on methods that are costly, slow, and technically complex, such as qPCR or next-generation sequencing. For the widescale screening of oral microorganisms suitable for point-of-care settings, there exists the need for a low-cost, rapid detection technique. Here, we tailored the novel CRISPR-Cas-based assay SHERLOCK for the species-specific detection of oral bacteria. We developed a computational pipeline capable of generating constructs suitable for SHERLOCK and experimentally validated the detection of seven oral bacteria. We achieved detection within the single-molecule range that remained specific in the presence of off-target DNA found within saliva. Further, we adapted the assay for detecting target sequences directly from unprocessed saliva samples. The results of our detection, when tested on 30 healthy human saliva samples, fully aligned with 16S rRNA sequencing. Looking forward, this method of detecting oral bacteria is highly scalable and can be easily optimized for implementation at point-of-care settings.

## Introduction

Oral polymicrobial dysbiosis has been well established as an etiology for numerous oral diseases, while a healthy oral microbiome has been suggested as crucial to preventing oral and systemic diseases [1–7]. In the case of periodontitis, a specific subgroup of microorganisms termed the red complex has been strongly associated with disease [1]. For caries, multiple acid-producing bacteria have been implicated in cavity formation [8,9]. Furthermore, specific bacteria such as *Streptococcus mutans*, *Porphyromonas gingivalis*, and *Fusobacterium nucleatum* are more intensely studied and have been highly associated with oral diseases [7,10–12]. These bacteria and others are also implicated in systemic diseases including various cancers, digestive diseases, cardiovascular diseases [3], and neurodegenerative diseases [13,14].

To detect microbes residing in the oral cavity, current methods frequently rely on quantitative polymerase chain reaction (qPCR) or next-generation sequencing (NGS). These techniques often require shipment to a central laboratory, high processing costs, and specialty technicians to interpret the data. Due to these limitations, despite the ubiquitous use of these techniques among the scientific community, large-scale and widespread employment of qPCR or sequencing-based methods for profiling oral microbes has remained limited [15,16]. To

facilitate precision oral medicine that can effectively treat conditions associated with oral bacteria, there is a clear need for the field of oral health to move towards novel diagnostics suitable for widespread use. To this end, we propose a recently developed detection method, Specific High-sensitivity Enzymatic Reporter Unlocking (SHERLOCK) [17,18], as a means to detect oral bacteria in a rapid, highly modular fashion.

SHERLOCK is a nucleic acid detection technique that leverages the CRISPR-Cas Class II enzyme Cas13a. It has been recently employed to detect a variety of pathogens including SAR-CoV-2 [19], various strains of ZIKA and Dengue virus [18], the malaria-causing parasites *P. falciparum* and *P. vivax* [20], and bacterial pathogens such as *Escherichia coli* and *Pseudomonas aeruginosa* [18]. SHERLOCK detection of nucleic acids consists of four key steps: (1) amplification of a target region; (2) Cas13a CRISPR-RNA (crRNA) recognition of the target region; (3) cleavage of reporter RNAs by Cas13a; and (4) fluorescence readout of cleaved reporter RNAs. The first step, amplification of a target region, utilizes the isothermal amplification technique Recombinase Polymerase Amplification (RPA) [21,22] to generate many copies of a target amplicon. Next, a crRNA, bound to Cas13a in a complex, recognizes a 28 base pair (bp) region on the target amplicon, a protospacer [18]. Upon crRNA binding to the protospacer, Cas13a exhibits indiscriminate RNase activity, cleaving nearby RNAs [18]. SHERLOCK

---

ⓘ Supplemental data for this article can be accessed online at https://doi.org/10.1080/20002297.2023.2207336.

leverages the collateral cleavage activity of Cas13a by including reporter RNAs within the reaction that fluoresce when cleaved [17]. This fluorescent signal can be quantified over time and equated to detection of the specific target sequence.

In combining Cas13a detection with RPA, SHERLOCK is capable of rapid detection in the single-molecule range without the need for complex equipment [12]. Furthermore, SHERLOCK reagents, including crRNAs and primers, can be lyophilized and employed on a laminar flow strip for an estimated cost of $0.60 per single-plex assay [17,18]. These improvements do not compromise the sensitivity of the assay compared to current gold-standard nucleic acid detection methods such as qPCR and ddPCR (droplet digital PCR) [18]. For these reasons, SHERLOCK and similar CRISPR-Cas-based assays have been proposed as ideal detection tools for future diagnostics [23].

To our knowledge, SHERLOCK has not yet been applied to the detection of oral bacteria. Here, we present the adaptation of SHERLOCK to detect oral bacteria by designing RPA primer pairs and crRNAs that target conserved, species-specific genes encoded within target bacteria. We experimentally validate detection using our designed constructs for seven oral bacteria and display that our crRNAs and primer pairs are highly specific and sensitive. We further tailor the SHERLOCK methodology to detect bacteria within unprocessed saliva samples. These results serve as a proof of principle that SHERLOCK can be a facile, rapid, and highly precise detection method for oral microorganisms with the potential to be further scaled and developed for use in point-of-care settings.

## Materials and methods

### Computational pipeline for species-specific primer and crRNA design

Given a bacterial species, our pipeline extracts all MetaPhlAn 4.0 marker genes attributed to that particular species [24]. The pipeline divides each marker gene into every possible 28 bp protospacer. For each possible protospacer, the pipeline searches for ancillary primers compatible with RPA using Primer3 [25] with the previously described thresholds of [17] primer length between 25 and 35 bp; Tm between 54 and 67°C; and GC% between 20 and 80%. To the 5' end of each left primer, a T7 polymerase promoter sequence is added (5'-AATTCTAATACGACTCACTATAGGGTCCA-3') [17]. crRNA sequences are then generated from each protospacer by adding a crRNA backbone sequence (5'-GGGGAUUUAGACUACCCCAAAAACGAAGGGG-ACUAAAAC-3') [18] to the 5' of the reverse complement of the protospacer.

For each candidate crRNA and its corresponding primer pair (hereafter referred to as a crRNA-primer set), the putative RNA amplicon produced by RPA and T7 transcription is predicted. The minimum free energy (MFE) of the RNA amplicon secondary structure is computed using ViennaRNA v2.5.1 [26]. As a higher secondary structure MFE of the crRNA target has been suggested to increase crRNA-protospacer binding [17,18], crRNA-primer sets are sorted in decreasing RPA amplicon secondary structure MFE. For the synthesis of crRNAs, a ssDNA template sequence is generated by adding a T7 promoter sequence to the 3' end of the reverse complement of the crRNA sequence (5'-TATAGTGAGTCGTATTAATTTC-3').

Using our computational pipeline, we synthesized seven crRNA-primer sets, each targeting one of seven oral bacteria (Table 1). All primers and crRNA-ssDNA templates were ordered from Integrated DNA Technologies (IDT). crRNAs were synthesized using a ssDNA template, a T7 primer (5' - GAAA TTAATACGACTCACTATAGGG – 3'), and the HiScribe T7 kit (NEB, #E2050S).

### Growth of target bacteria and gDNA isolation

Seven target bacteria were grown in the conditions listed in Table 1. Genomic DNA (gDNA) of each target bacterium was isolated using a previously described method [41]. Briefly, the MasterPure gram-positive DNA purification kit (Biosearch, #MC85200) was used according to the manufacturer's protocol with the addition of a bead beating step ($3 \times 30$ s at 6 m/s with 1-min pause intervals). Isolated gDNA concentrations were measured using the Qubit Broad Range dsDNA kit (Invitrogen, #Q32850). In

**Table 1.** Bacteria targeted in this study and their disease association, genome size, and growth conditions.

| Bacterial Species | Strain Name | Disease Association | Genome Size (bp) | Growth Conditions | References |
|---|---|---|---|---|---|
| *Streptococcus mutans* (Sm) | F0577 | Caries | 2030936 | Brain Heart Infusion Medium 23% $O_2$, 77% $N_2$ | [27,28] |
| *Scardovia wiggsiae* (Sw) | F0424 | Caries | 1550000 | Brain Heart Infusion Medium 0% $O_2$, 5% $CO_2$, 95% $N_2$ | [8,29] |
| *Aggregatibacter actinomycetemcomitans* (Aa) | AA075 | Periodontitis | 2260000 | Brain Heart Infusion Medium 23% $O_2$, 77% $N_2$ | [30,31] |
| *Porphyromonas gingivalis* (Pg) | ATCC33277 | Periodontitis | 2378872 | Brain Heart Infusion Medium 0% $O_2$, 5% $CO_2$, 95% $N_2$ | [32,33] |
| *Acinetobacter baumannii* (Ab) | 19606 | Oral and Systemic Disease | 3824000 | Brain Heart Infusion Medium 23% $O_2$, 77% $N_2$ | [34,35] |
| *Klebsiella pneumoniae* (Kp) | Fpn6806 | Oral and Systemic Disease | 5491870 | Brain Heart Infusion Medium 0% $O_2$, 5% $CO_2$, 95% $N_2$ | [36–38] |
| *Staphylococcus aureus* (Sa) | F0253A | Oral and Systemic Disease | 2874302 | Brain Heart Infusion Medium 23% $O_2$, 77% $N_2$ | [34,39,40] |

the case of isolated gDNA from axenic cultures, gDNA concentration and copy number/uL were calculated based on established genome lengths (Table 1).

### One-pot SHERLOCK

SHERLOCK detection was performed in a one-pot reaction, according to a previously described protocol [17], in which the *Leptotrichia wadei* ortholog of Cas13a (*lw*Cas13a), RPA reagents, primers, crRNA, and several adjuvant reagents are combined. Reagents were combined in the specific order and at volumes listed in Appx. Table A1 then pulse mixed for 3 s. RPA reagents from the TwistAmp Basic Kit (TwistDx, #TABAS03KIT) were used. A Spectra Max iD3 Multi-Mode microplate reader was preheated to 37°C. Target gDNA was diluted to the desired concentrations and was added at a volume of 4 μL into 20.6 μL of the Cas-RPA mix. The final reaction mixture was homogenized and centrifuged at 500 $g$ for 30 s before 20 μL was pipetted into a 384-well optical plate (Roche, #5102430001). The plate was loaded into the microplate reader and fluorescence readings of each well were taken at 485/528 nm using kinetic reads at 5-min intervals.

### Bacterial detection directly from unprocessed saliva samples

To adapt the SHERLOCK system for the detection of bacteria in unprocessed saliva samples (without a DNA extraction step), we modified a previously described SHERLOCK protocol [19] in which EGTA (ethylene glycol-bis(beta-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid) and DTT (Dithiothreitol) are added to a saliva sample. Briefly, ETGA and DTT were added at a concentration of 500 mM and 1 mM, respectively, to an unprocessed saliva sample. The saliva mixture was then heated at 95°C for 15 min. The saliva mixture, in lieu of target gDNA, was then used in a one-pot SHERLOCK assay (this methodology is hereafter referred to as SHERLOCK-EGTA+DTT).

In our spiked saliva experiments, to determine the amount of spiked target bacteria, we estimated the CFU/μL of live bacteria using previously established optical density to CFU measurements for both Pg [42] and Sm [43]. Unprocessed saliva samples were collected from medically healthy subjects with no oral diseases or conditions (see below).

### Clinical sample collection and validation

We collected unstimulated saliva samples (1.5 mL) from 30 adults (18 females, 12 males) with documented medical, dental, and periodontal health. As previously described [44], participants were excluded if they received periodontal disease treatment within 3 months, antibiotic treatment within 8 weeks, or antimicrobial mouthwash treatment within 4 weeks. Additionally, participants with less than 10 teeth or those currently using dentures were excluded. The study received Forsyth Institutional Review Board approval (#18–06, Dr Hasturk), and all subjects provided written informed consent. Saliva samples were split into two aliquots: one for 16S rRNA sequencing and one for SHERLOCK-EGTA+DTT detection. gDNA was isolated from an aliquot of each saliva sample, and V3-V4 16S rRNA sequencing was performed on all 30 samples (Zymogen). 16S rRNA taxonomic assignments and relative abundances measurements were determined by comparing in-house Zymogen analyses with 16S rRNA sequences from the expanded Human Oral Microbiome Database (eHOMD) [45]. Briefly, the in-house Zymogen analysis utilizes DADA2 [46] to determine amplicon sequence variants (ASVs) and ASV relative abundances within each sample. The Zymogen generated ASVs were then compared to the eHOMD 16S rRNA database for taxonomic assignment as previously described [41]. To generate *Streptococcus* and *Porphyromonas* phylogenetic trees, *Streptococcus* and *Porphyromonas* 16S rRNA sequences were downloaded from eHOMD. One 16S rRNA sequence was selected for each Human Microbial Taxon (HMT), and the 16S rRNA sequence of *Escherichia coli* from eHOMD was included as an outgroup. Sequences were aligned using MAFFT [47] and trimmed using TrimAl [48]. Maximum-likelihood trees were inferred using IQ-tree v2.1.4-beta with ultrafast bootstrap (–bb 1500) [49]. For SHERLOCK-EGTA+DTT detection, the second aliquot of each saliva sample was diluted 1:4 with sterile PBS before addition to SHERLOCK reagents.

## Results

### Computational pipeline generates constructs compatible with SHERLOCK

To detect a specific nucleic acid sequence, SHERLOCK (Figure 1a, Appx. Table A1) requires a set of three constructs specific to the target sequence (hereafter referred to as a crRNA-primer set): a crRNA, a forward primer for RPA, and a reverse primer for RPA. To design crRNA-primer sets for the SHERLOCK detection of oral bacteria, we compiled a computational pipeline (Figure 1b) that generates crRNA-primer sets targeting conserved species-specific genes (Figure 1c, *Materials and Methods*).

To experimentally validate our computational pipeline, we synthesized seven crRNA-primer sets, each targeting a species-specific gene encoded within one of seven oral bacteria. We targeted *Streptococcus mutans* (Sm), *Porphyromonas gingivalis* (Pg), *Scardovia wiggsiae* (Sw), *Aggregatibacter actinomycetemcomitans* (Aa), *Klebsiella pneumoniae* (Kp), *Acinetobacter baumannii* (Ab), and *Staphylococcus aureus* (Sa) for detection (Table 1, Appx. Table A2). As an assessment of the seven crRNA-primer sets, we tested each crRNA-primer
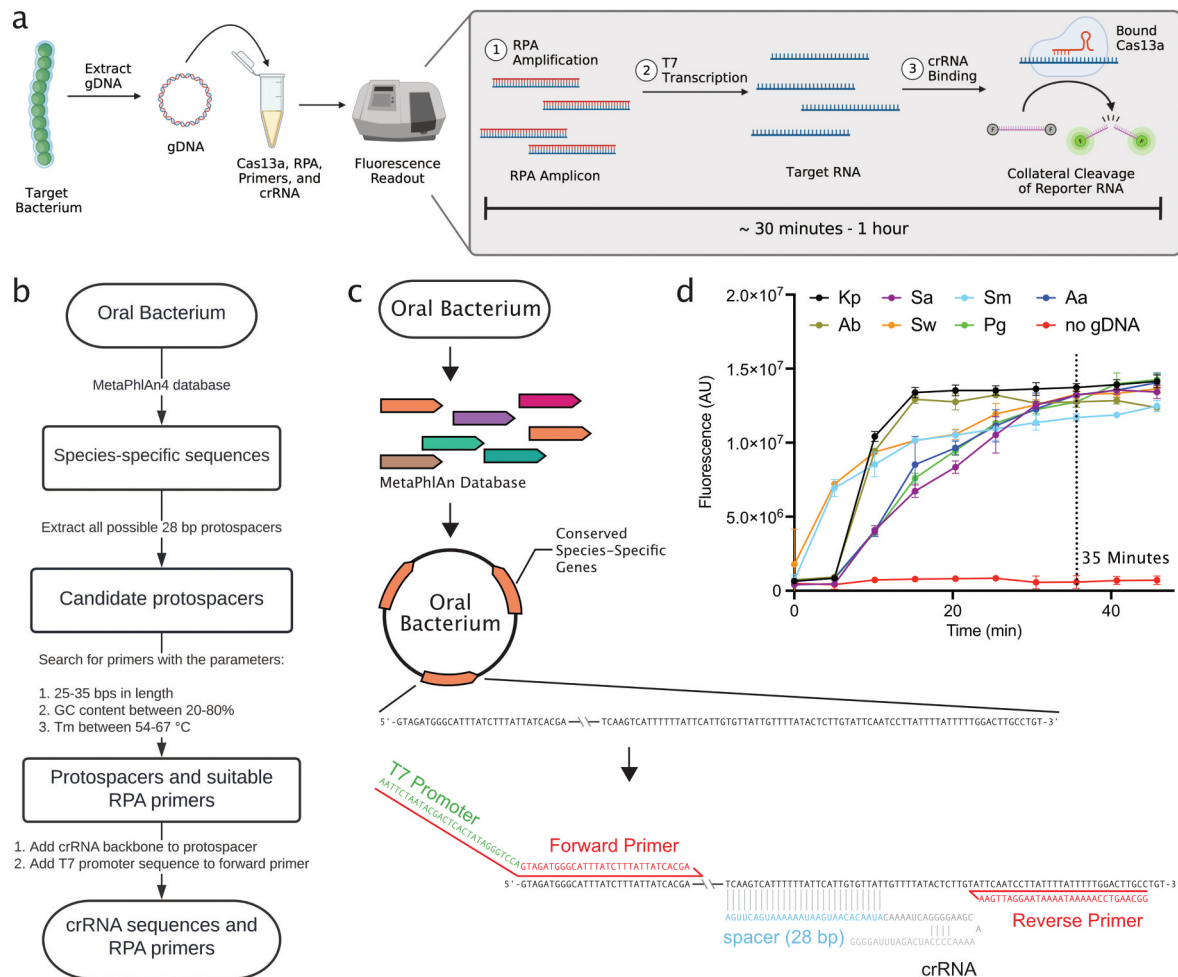
**Figure 1.** Computational pipeline to generate crRnas and primer pairs targeting conserved species-specific genes. a. Visual schematic of one-pot SHERLOCK (Appx Table 1). b. Workflow of the computational pipeline that generates crRnas and primers for specific bacterial detection using SHERLOCK. c. Schematic illustrating the architecture of crRNA and primer design for SHERLOCK. d. Detection of target gDNA using each synthesized crRNA. Target gDNA was added at a concentration of 1 picomolar (pM) per reaction, equating to 30,000 targets per μL reaction. Each crRNA targeted one of seven oral bacterial targets: *Streptococcus mutans* (Sm), *Klebsiella pneumoniae* (Kp), *Aggregatibacter actinomycetemcomitans* (Aa), *Acinetobacter baumannii* (Ab), *Staphylococcus aureus* (Sa), *Scardovia wiggsiae* (Sw), or *Porphyromonas gingivalis* (Pg).

set with SHERLOCK in the presence of genomic DNA (gDNA) isolated from their intended bacterial target. All seven crRNA-primer sets produced a target-specific fluorescent signal over time. At 35 min of detection, the typical time at which the fluorescent signals plateaued, the fluorescent signals were greater than 19-fold change in the presence of target gDNA (1 picomolar (pM); 30000 copies/μL reaction) as compared to a negative control containing no target gDNA (Figure 1d). These results confirmed that our computational pipeline produced crRNA-primer sets compatible with SHERLOCK detection.
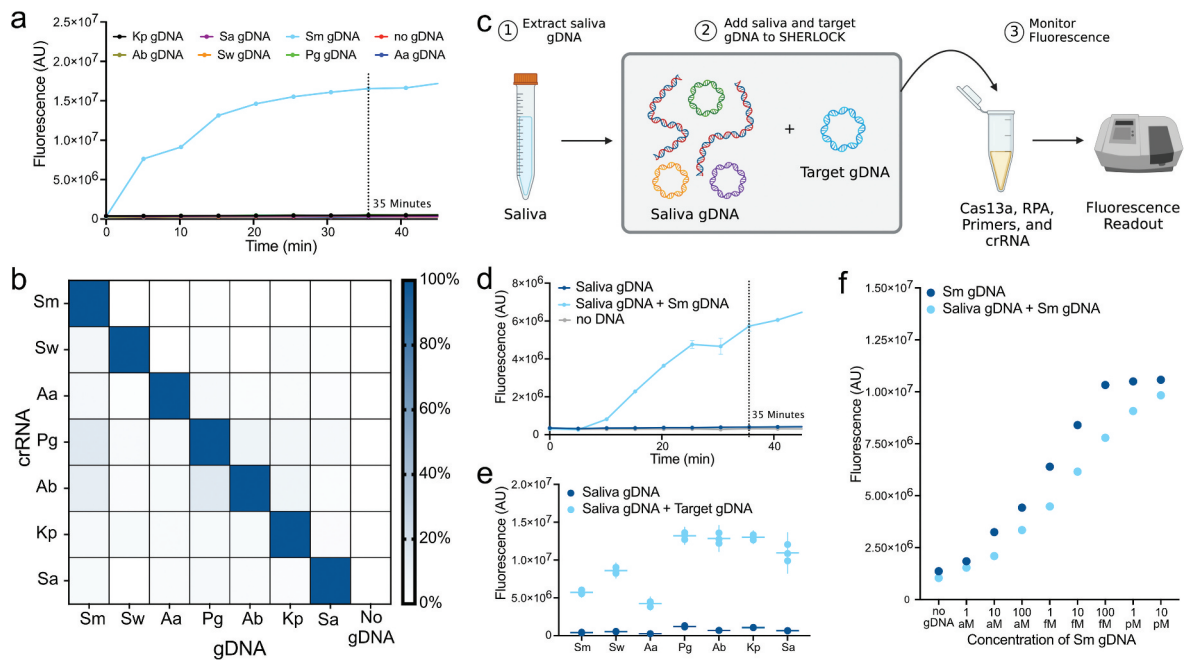
## Designed crRNAs display specific and sensitive signal

With our seven synthesized crRNA-primer sets displaying signal in the presence of their target sequence, we next sought to examine the specificity and sensitivity of our designed crRNA-primer sets. As a preliminary measure

of the specificity of the designed crRNA-primer sets, we tested the cross-specificity of each crRNA-primer set against isolated gDNA (1 pM) from each of the seven bacterial targets (Figure 2a). At 35 min of detection, when in the presence of gDNA isolated from their intended target, all crRNA-primer sets produced a fluorescent signal greater than 10-fold change higher than off-target reactions (Figure 2b, Appx. Figure A1). These cross-specificity assays indicated that when tested against gDNA from several off-target oral bacteria, all designed crRNA-primer sets produced a clear and discriminant signal only against their intended target.

Next, to account for the hundreds of endogenous species within saliva, we conducted a more thorough interrogation of the specificity of each crRNA-primer set by testing them against a background of saliva gDNA (Figure 2c). We surveyed the fluorescence signal at 35 min of detection produced by each crRNA-primer set when tested against (1) saliva gDNA (10 ng); (2) saliva gDNA (10 ng) spiked with target bacterial gDNA (1 pM);

**Figure 2.** Specificity and sensitivity of designed crRNA and primer pairs. a. Representative fluorescent trace plot displays the specificity of the Sm crRNA-primer set against other bacterial targets. b. Heat map displaying cross-reactivity of synthesized crRNA-primer sets against all tested oral bacterial targets. The min-max linear normalized signal at 35 min of detection is displayed. Target gDNA was added at a concentration of 1 pM per reaction. All other corresponding individual plots can be found in Appx. Fig. 1. c. Visual schematic displaying specificity assays with saliva gDNA. d. Representative example fluorescence trace plot displays the specificity of the Sm crRNA-primer set against background saliva gDNA. Other individual plots can be found in Appx. Fig. 2. e. the specificity of synthesized crRNA-primer sets against saliva gDNA. Signal at 35 min of detection is displayed. Each point represents a technical replicate. 10 ng of background saliva gDNA was added per reaction. Target gDNA was added at a concentration of 1 pM per reaction. f. SHERLOCK detection at 35 min of decreasing 10-fold concentrations of Sm gDNA alone or combined with 10 ng of saliva gDNA (Appx. Fig. 3).

and (3) no saliva or target bacterial gDNA (Figure 2d). Each crRNA-primer set displayed a greater than 12-fold change increase in fluorescent signal when in the presence of gDNA from their bacterial target as compared to saliva gDNA alone (Figure 2e; Appx. Figure A2). Intriguingly, at 35 min of detection, the different crRNA-primer sets displayed differing levels of fluorescent signal, indicating that they may express varying levels of efficiency (Figure 2e). Altogether, these results validated the specificity of our crRNA-primer sets in the presence of off-target oral bacterial and human gDNA.

The addition of off-target saliva gDNA may also affect the sensitivity of SHERLOCK detection. To assess this, we performed SHERLOCK detection against a concentration gradient of Sm gDNA ranging from 1 attomolar (aM) (0.5 copies/μL reaction) to 10 pM (300,000 copies/μL reaction) in two conditions: (1) with 10 ng of background saliva gDNA and (2) without any background gDNA. Across the Sm gDNA dilutions, the saliva gDNA background dampened the fluorescent signal strength by an average of 20% at 35 min of detection (Figure 2f, Appx. Figure A3). Despite this decrease in signal strength incurred by the addition of saliva gDNA, at our lowest concentration of Sm gDNA, 1 aM (0.5 copies/μL reaction), we still achieved a fluorescent signal 1.4-fold change greater with Sm gDNA than without Sm gDNA in a

background of saliva gDNA (Appx. Figure 3). The ability of our assay, therefore, to detect target sequences down to a concentration of 1 aM was not hindered by the addition of a saliva gDNA background. Further, in assessing the detection signal at 35 min across dilutions of target gDNA, there is a clear correlation between target DNA concentration and measured signal (Figure 2f).

## Detection of bacteria from unprocessed saliva

With our synthesized crRNA-primer sets displaying high specificity and sensitivity, we modified a previously described protocol for SHERLOCK on unprocessed saliva samples [19] to extend the utility of our bacterial detection assay (Figure 3a, *Materials and Methods*). EGTA and DTT were added to inhibit the saliva RNases while a 95°C treatment was added to lyse bacteria. We hereafter refer to this enhanced method as SHERLOCK-EGTA+DTT.

We sought to assess how the SHERLOCK-EGTA +DTT detection of live bacteria in unprocessed saliva samples compared to SHERLOCK detection of isolated gDNA. We chose to compare the two methodologies using gram-negative Pg and gram-positive Sm. At 35 min of Sm or Pg detection, when equivalent amounts of cfu/μL and gDNA copy numbers
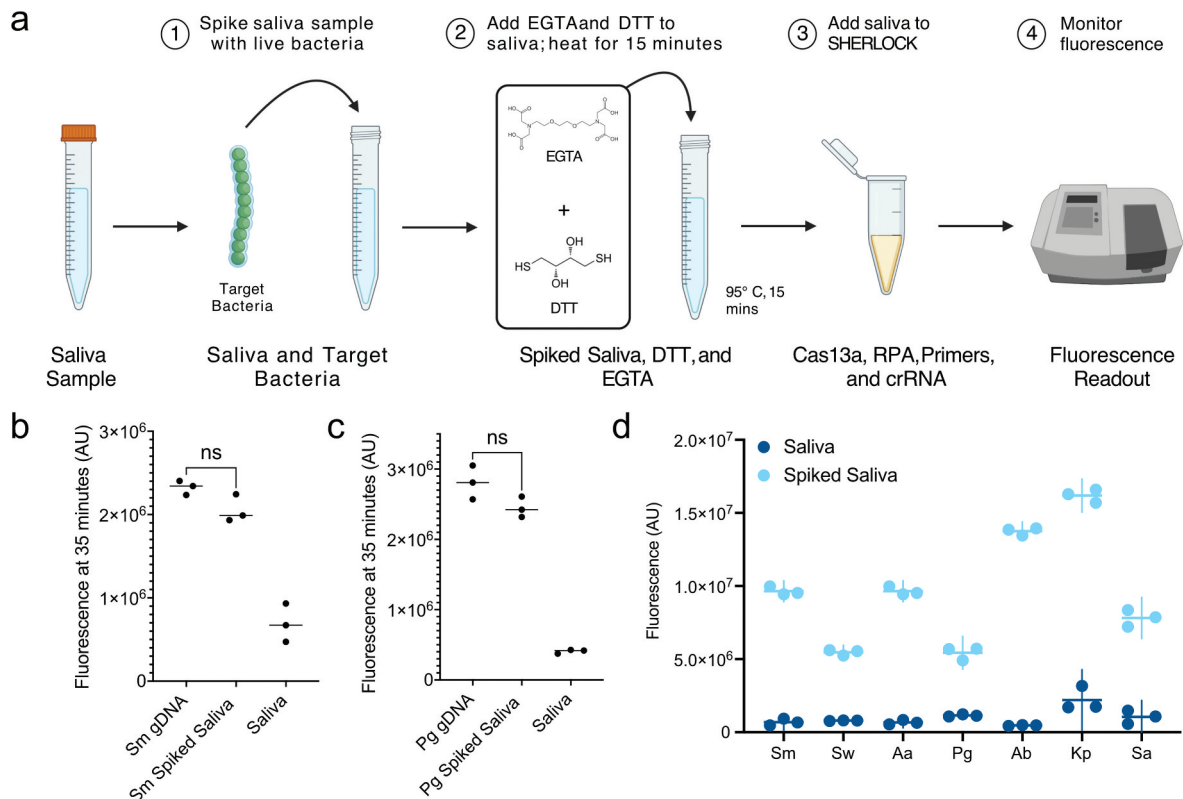
**Figure 3.** Oral pathogen detection directly from saliva samples. a. Visual schematic of unprocessed saliva detection with SHERLOCK-EGTA+DTT. b-c. Fluorescent signal at 35 min produced by SHERLOCK with Sm (B) or Pg (C) gDNA and SHERLOCK-EGTA+DTT spiked with an equivalent concentration of live Sm or Pg cells (Appx. Fig. 4). d. Specificity and fluorescent signal of crRNA-primer sets when used with SHERLOCK-EGTA+DTT on unprocessed saliva samples. Saliva was spiked at a concentration of 300 target bacterial cells per μL reaction. Fluorescent signal at 35 min is displayed. Each point represents a technical replicate. All plots can be found in Appx. Fig. 5.

(300 cfu/μL reaction; 300 copies/μL reaction) were added to SHERLOCK-ETGA+DTT and SHERLOCK, respectively, we did not observe a significant difference between the fluorescent signals produced by the two methodologies (Figure 3b,c, Appx. Figure A4). Therefore, in the case of Sm and Pg detection, the SHERLOCK-EGTA+DTT methodology is as effective as SHERLOCK is at detecting target nucleic acid sequences from isolated gDNA.

We tested the ability of each of the seven synthesized cRNA-primer sets using SHERLOCK-EGTA +DTT to detect their target bacterium from unprocessed saliva. All seven crRNA-primer sets with SHERLOCK-EGTA+DTT produced a fluorescent signal greater than 6-fold change higher when tested against unprocessed saliva spiked with live target bacterium (300 cfu/μL reaction) compared to unprocessed saliva without the target bacteria (Figure 3d, Appx. Figure A5). As observed in our isolated gDNA detection (Figure 2e), different crRNA-primer sets displayed varying levels of fluorescent signal at 35 min of detection (Figure 3d). However, this did not affect detection results. In no reaction did unprocessed saliva samples without spiked target bacteria produce a fluorescent signal greater than 19% of

samples with the target bacteria (Appx. Figure A5). These two observations indicated that our added reagents and additional heating step in SHERLOCK-EGTA+DTT effectively neutralized endogenous RNase activity and lysed bacterial cells. Additionally, these data indicate that our designed crRNA-primers sets remained effective when used for detection from unprocessed saliva samples.

## Detection of target bacteria in unprocessed human saliva aligns with 16S rRNA sequence profiling

To be effective in healthcare and research settings, it is important that a novel detection method is similar in sensitivity and accuracy to established methods such as 16S rRNA sequencing [15,16,50]. To demonstrate that SHERLOCK-EGTA+DTT detection is similar in detection sensitivity to 16S rRNA sequencing, we collected 30 saliva samples from healthy subjects and performed 16S rRNA sequencing on each sample. Within the 30 saliva samples, 16S rRNA sequencing identified 202 different oral bacterial species (Figure 4a). These saliva samples each contained between 33 and 102 oral bacterial species,
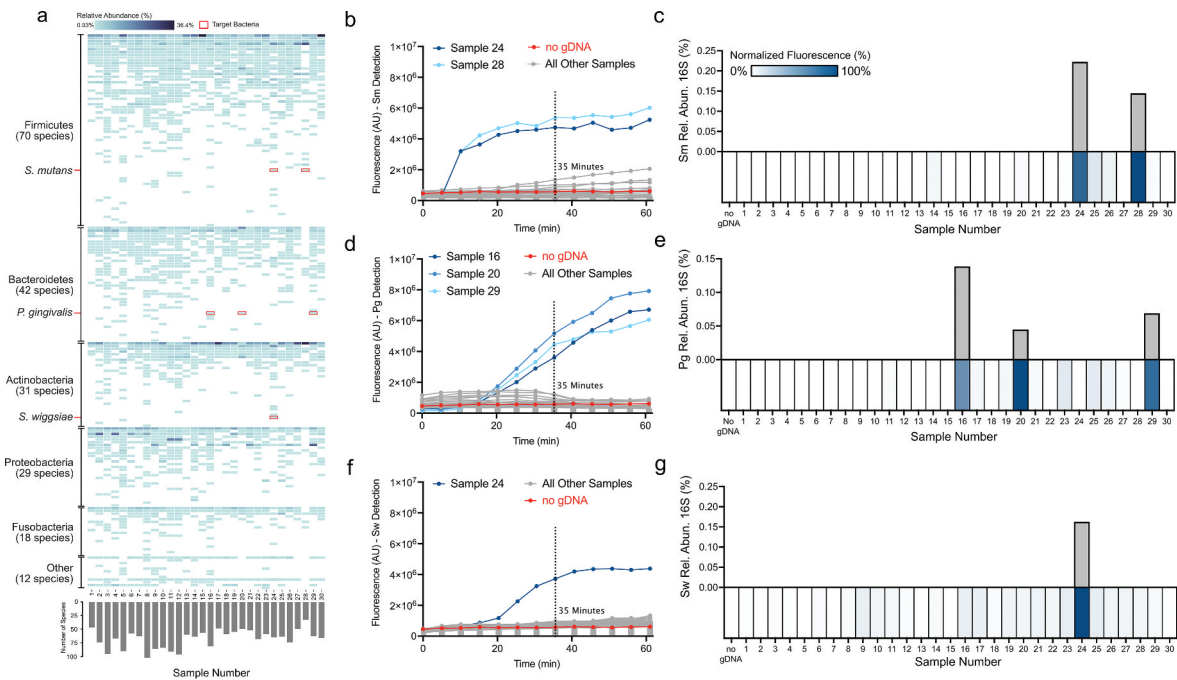
**Figure 4.** 16S rRNA sequencing validation of SHERLOCK-EGTA+DTT detection on healthy subject saliva. a. 16S rRNA taxonomic predictions for 30 healthy saliva samples. b. SHERLOCK-EGTA+DTT detection of 30 healthy subject samples using the Sm crRNA-primer set. Six microlitres of 1:4 diluted healthy subject saliva was added per reaction. c. 16S rRNA sequencing compared to SHERLOCK-EGTA+DTT Sm detection. Upper bar graph displays 16S rRNA relative abundance (%) of Sm. Lower heatmap displays the min-max linear normalized signal of Sm SHERLOCK-EGTA+DTT detection at 35 min. d. SHERLOCK-EGTA+DTT detection of 30 healthy subject samples using the Pg crRNA-primer set. Six microlitres of 1:4 diluted healthy subject saliva were added per reaction. e. 16S rRNA sequencing compared to SHERLOCK-EGTA+DTT Pg detection. Upper bar graph displays 16S rRNA relative abundance (%) of Pg. Bottom heatmap displays the min-max linear normalized signal of Pg SHERLOCK-EGTA+DTT detection at 35 min. f. SHERLOCK-EGTA+DTT detection of 30 healthy subject samples using the Sw crRNA-primer set. Six microlitres of 1:4 diluted healthy subject saliva were added per reaction. g. 16S rRNA sequencing compared to SHERLOCK-EGTA+DTT Sw detection. Upper bar graph displays 16S rRNA relative abundance (%) of Sw. Bottom heatmap displays the min-max linear normalized signal of Sw SHERLOCK-EGTA+DTT detection at 35 min.

a typical range found in other studies [51]. Of our seven candidate oral bacteria, 16S rRNA sequencing identified two (samples 24 and 28), three (samples 16, 20, and 29), and one (sample 24) samples positive for Sm, Pg, and Sw, respectively (Figure 4a). None of the other four bacterial targets (Aa, Kp, Sa, and Ab) were identified in any of the 30 samples by 16S rRNA sequencing. This was not surprising given that these bacteria are rarely found in healthy oral cavities [52].

We performed SHERLOCK-ETGA+DTT on all 30 samples using the Sm, Pg, and Sw crRNA-primer sets. By Sm detection, samples 24 and 28 produced a signal greater than 8-fold change over a no gDNA control, while no other samples produced a signal greater than 1.8-fold change over the negative control (Figure 4b). These results completely aligned with the 16S rRNA detection of Sm (Figure 4c). Similarly, by Pg detection, we observed only three strong fluorescent signals (all greater than sixfold change over the negative control) (Figure 4d). The three signals were from the same samples that were identified to contain Pg by 16S rRNA sequencing (Figure 4e). With Pg crRNA, we also observed the lowest relative abundance by 16S rRNA sequencing, 0.05%, was discernible by SHERLOCK-EGTA+DTT, illustrating the

high sensitivity of the assay. By Sw detection, we observed only one discernible signal: the signal produced by sample 24 was 6.5-fold change greater than the negative control (Figure 4f). This was the same sample identified to contain Sw by 16S rRNA sequencing (Figure 4g). Altogether, our SHERLOCK-ETGA +DTT detection results accurately recapitulated the 16S rRNA sequencing results for Sm, Pg, and Sw presence (100% agreement between the two methods).

To further assess the species-specificity of our designed crRNAs, we surveyed the community composition of the 30 saliva samples and detected multiple species of *Streptococcus* and *Porphyromonas* (Figure 5). Of the 37 *Streptococcus* oral Human Microbial Taxons (HMTs – defined as 98.5% sequence identity between 16S rRNA genes) within the expanded Human Oral Microbiome Database (eHOMD) [45], we detected 17 *Streptococcus* HMTs within our 30 saliva samples (Figure 5a). The 17 detected *Streptococcus* HMTs are not confined to a specific clade and capture a large range of the oral *Streptococcus* genus. As evidence of the species specificity of our designed crRNAs, only samples containing *S. mutans*
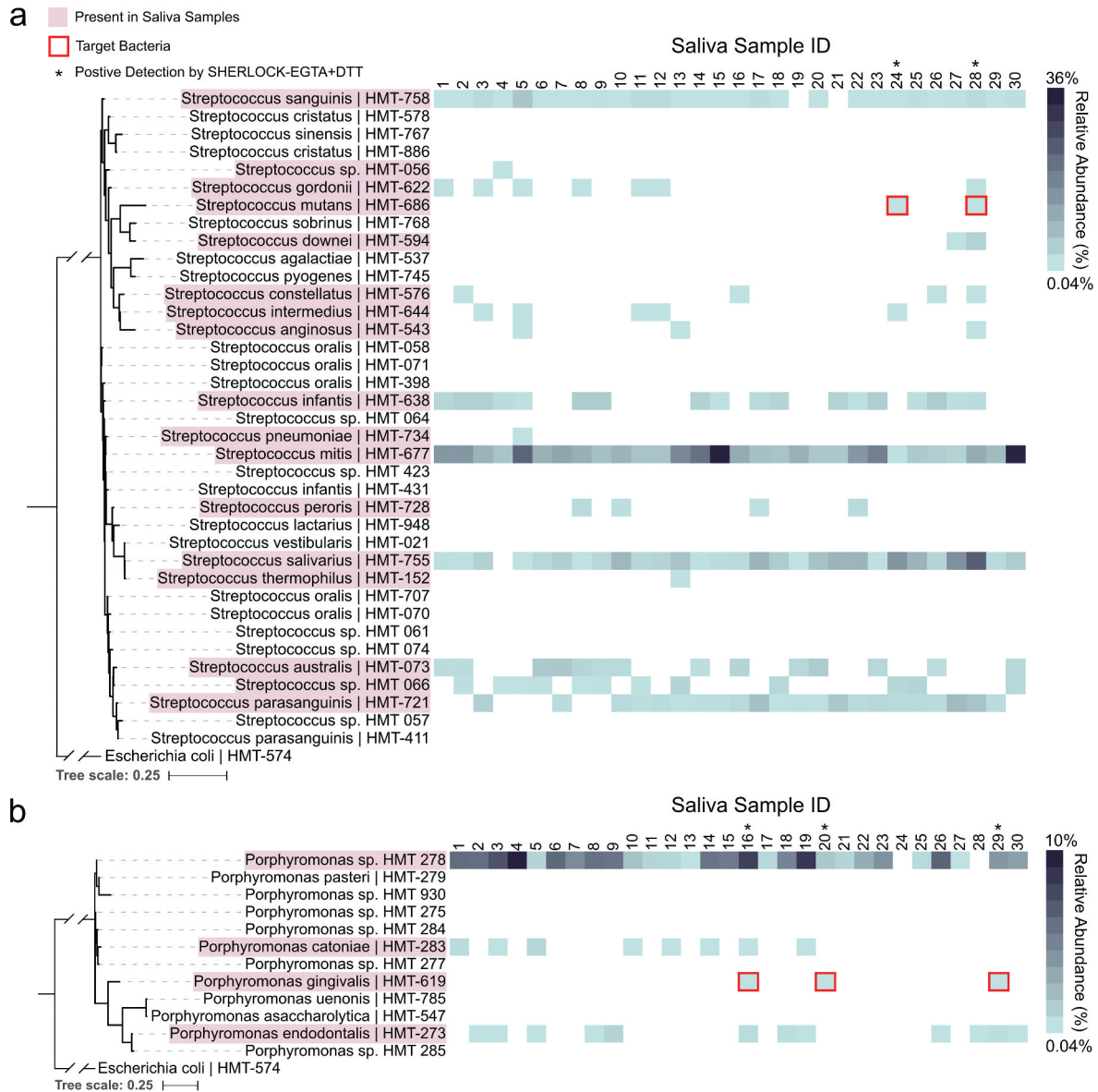
Figure 5. Species specificity of *S. mutans* and *P. gingivalis* detection. Pink outlines denote HMTs that were detected within our 30 saliva samples. Blue heatmaps display the relative abundance of the detected HMTs within each of the 30 saliva samples. a. Maximum-likelihood 16S rRNA phylogenetic tree of eHOMD *Streptococcus* HMTs. Asterisk denotes samples that displayed fluorescent signal in Fig. 4b,c when tested with SHERLOCK-EGTA+DTT and the designed Sm crRNA. *E. coli* is included as an outgroup. Branches leading to *E. coli* and the *Streptococcus* genus are truncated to improve readability. b. Maximum-likelihood 16S rRNA phylogenetic tree of eHOMD *Porphyromonas* HMTs. Asterisk denotes samples that displayed fluorescent signal in Fig. 4d,e when tested with SHERLOCK-EGTA+DTT and the designed Pg crRNA. *E. coli* is included as an outgroup. Branches leading to *E. coli* and the *Porphyromonas* genus are truncated to improve readability.

(samples 24 and 28) displayed a fluorescent signal when tested using SHERLOCK-EGTA+DTT with our Sm crRNA despite the presence of many *Streptococcus* HMTs across the thirty saliva samples (Figure 4b and c). Samples containing closely related *Streptococcus* species such as *S. downei* but not *S. mutans* (sample 27) did not exhibit detection by our Sm crRNA. As additional evidence, our designed Pg crRNA displayed similar specificity. Of the 12 eHOMD *Porphyromonas* HMTs, we detected four within our saliva samples (Figure 5b). Only the samples containing *P. gingivalis* (samples 16, 20, and 29) exhibited a fluorescent signal when tested with SHERLOCK-EGTA+DTT

and our Pg crRNA (Figure 4d and e). Although we did not test the species-specificity of the other five crRNAs, these two examples demonstrate that our designed crRNAs, in the case of Sm and Pg detection and with respect to the species we identified within our saliva samples, are indeed species-specific.

## Discussion

While SHERLOCK has been used previously to detect pathogens and parasites, this study marks the first time it has been applied to detect oral bacteria. We designed a computational pipeline capable of

generating species-specific crRNAs and primer pairs for SHERLOCK and demonstrated highly specific and sensitive detection for seven oral bacteria. Further, we tailored SHERLOCK for the detection of oral bacteria directly from unprocessed saliva samples and validated this enhanced method using healthy oral saliva samples. Our results suggest that crRNA-primer sets paired with SHERLOCK-EGTA +DTT are on par with canonical oral microbial detection techniques such as 16S rRNA sequencing. Alternative oral bacterial detection technologies and services that come closest to point-of-care settings are, to our knowledge, OralDisk [53], OralDNA Labs [54], and Bristle [16]. However, these approaches are based on either qPCR or NGS techniques. They therefore require complicated hardware, significant amounts of reagent, and up to 7 days of processing time. We envision that our assay can be developed for similar applications – assessing oral microbial profiles in dental clinics to assist dentists in the prevention and treatment of oral dysbiotic diseases – and could be utilized in a much more rapid manner with less labor required. For further ease of use, SHERLOCK has been previously employed on laminar flow strips and may be similarly applied for the detection of oral bacteria [17].

To accomplish the detection of oral bacteria using SHERLOCK, we solved multiple challenges specific to oral saliva samples such as computationally designing crRNAs that target oral bacteria, preparing saliva samples for SHERLOCK detection without a DNA isolation step, and testing the specificity of our crRNAs against a background of gDNA from human cells and hundreds of oral microbes. Although saliva contains numerous compounds that could interfere with SHERLOCK detection, such as endogenous RNases, off-target nucleic acid sequences, and salivary mucins, we did not observe any interference in our results using the SHERLOCK-EGTA+DTT methodology compared to SHERLOCK on isolated gDNA (Figure 3b and c). Previously, variations of this methodology have been combined with SHERLOCK or other Cas-based techniques to detect viruses and eukaryotic parasites in saliva [19,55,56], but they have not yet been applied to the detection of bacteria in saliva. Going forward, the addition of EGTA and DTT or similar adjuvant reagents to saliva samples may be critical to the future application of SHERLOCK and other nucleic acid-based detection systems to the oral field.

Limitations of our study included the testing of clinically healthy subjects without comparison to diseased patients. For this study, which was focused on establishing sensitive and specific detection, we reasoned that using healthy subject samples was sufficient. In addition, we observed that not all crRNA-primer sets displayed the same level of fluorescence when in the presence of equal amounts of target gDNA or live bacteria. These data suggest that in the case of gDNA detection, some crRNA-primer sets may display a higher efficiency than others. This could be due to RPA amplification, T7 transcription, crRNA detection, or some combination of the three steps. Further optimization of our crRNA-primer sets may prove fruitful in achieving higher fluorescent signals and faster detection times. Future studies will focus on testing additional crRNA-primer sets for each bacterial target. In addition, alternative isothermal amplification techniques to RPA, such as loop-mediated isothermal amplification [57] may display higher efficiencies and specificity.

Human oral diseases that have an etiology in microorganisms are generally accepted as polymicrobial in nature. Commensal microorganisms naturally present in the oral cavity, for instance, may become more abundant leading to dysbiosis and disease. Specific bacteria such as *P. gingivalis* and *S. mutans* are strongly associated with periodontitis and caries and are typically considered pathogenic [27,58]. Therefore, it is important to detect individual opportunistic pathogens and to profile the overall community. Within this scope, we have clearly illustrated in this study that SHERLOCK has the capability to detect specific oral bacteria and, with further development, may be capable of quantitating abundance (Figure 2f). Further, previous efforts to multiplex SHERLOCK for the detection of multiple organisms simultaneously have been successful [17,59]. The CARMEN system, for instance, is capable of simultaneously detecting 100+ human viruses [59]. Although not shown in our study, SHERLOCK detection has the potential to be expanded to other microorganisms in the oral cavity including viruses, fungi, and additional bacteria. Looking forward, SHERLOCK has the necessary capabilities (detection, quantitation, and multiplexing) to allow for the rapid profiling of an individual's oral microbiome. Our current findings provide a foundation for the use of SHERLOCK and possibly other CRISPR-Cas based detection technologies in the oral research and clinical fields and are an important step forward toward individualized and precise oral care.

## Data availability statement

All nucleic acid sequences are provided in the manuscript or in Appx. Table A2. Raw data for Figures 1d-A4 and all code used in this project are available on Zenodo (DOI: 10.5281/ zenodo.7708546). The raw data and code are also available at https://www.borlab.org/resources. Bacterial strains used in this paper will be provided upon request.

## ORCID

Batbileg Bor   http://orcid.org/0000-0002-1797-1730

## References

[1] Griffen AL, Beall CJ, Campbell JH, et al. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. Isme J. 2012;6 (6):1176–1185. DOI:10.1038/ismej.2011.191

[2] Kumar PS, Griffen AL, Barton JA, et al. New bacterial species associated with chronic periodontitis. J Dent Res. 2003;82(5):338–344. DOI:10.1177/154405910308200503

[3] Li X, Kolltveit KM, Tronstad L, et al. Systemic diseases caused by oral infection. Clin Microbiol Rev. 2000;13:547–558.

[4] Freire M, Nelson KE, Edlund A. The oral host–microbial interactome: an ecological chronometer of health? Trends Microbiol. 2021;29:551–561.

[5] Wade WG. The oral microbiome in health and disease. Pharmacol Res. 2013;69:137–143.

[6] Abusleme L, Dupuy AK, Dutzan N, et al. The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. Isme J. 2013;7(5):1016–1025. DOI:10.1038/ismej.2012.174

[7] Tanner ACR, Kressirer CA, Rothmiller S, et al. The caries microbiome: implications for reversing dysbiosis. Adv Dent Res. 2018;29:78–85.

[8] Kressirer CA, Smith DJ, King WF, et al. Scardovia wiggsiae and its potential role as a caries pathogen. J Oral Biosci. 2017;59(3):135–141. DOI:10.1016/j.job.2017.05.002

[9] Gross EL, Leys EJ, Gasparovich SR, et al. Bacterial 16S sequence analysis of severe caries in young permanent teeth. J Clin Microbiol. 2010;48(11):4121–4128. DOI:10.1128/JCM.01232-10

[10] Hajishengallis G. Immunomicrobial pathogenesis of periodontitis: keystones, pathobionts, and host response. Trends Immunol. 2014;35:3–11.

[11] Chen Y, Huang Z, Tang Z, et al. More than just a periodontal pathogen –the research progress on Fusobacterium nucleatum. Front Cell Infect Microbiol. 2022;12:815318.

[12] Han YW. Fusobacterium nucleatum: a commensal-turned pathogen. Curr Opin Microbiol. 2015;23:141–147.

[13] Dominy SS, Lynch C, Ermini F, et al. Porphyromonas gingivalis in Alzheimer's disease brains: evidence for disease causation and treatment with small-molecule inhibitors. Sci Adv. 2019;5(1):eaau3333. DOI:10.1126/sciadv.aau3333

[14] Wu H, Qiu W, Zhu X, et al. The periodontal pathogen Fusobacterium nucleatum exacerbates alzheimer's pathogenesis via specific pathways. Front Aging Neurosci. 2022;14:912709.

[15] Oravital home - Preventing oral systemic diseases. Available from: https://www.oravital.com/.

[16] Bristle: oral health test. Available from: https://www.bristlehealth.com/product-oral-health-test.

[17] Kellner MJ, Koob JG, Gootenberg JS, et al. SHERLOCK: nucleic acid detection with CRISPR nucleases. Nat Protoc. 2019;14(10):2986–3012.

[18] Gootenberg JS, Abudayyeh OO, Lee JW, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. Science. 2017;356(6336):438–442. DOI:10.1126/science.aam9321

[19] de Puig H, Lee RA, Najjar D, et al. Minimally instrumented SHERLOCK (miSHERLOCK) for CRISPR-based point-of-care diagnosis of SARS-CoV-2 and emerging variants. Sci Adv. 2021;7(32):eabh2944. DOI:10.1126/sciadv.abh2944

[20] Cunningham CH, Hennelly CM, Lin JT, et al. A novel CRISPR-based malaria diagnostic capable of Plasmodium detection, species differentiation, and drug-resistance genotyping. EBioMedicine. 2021;68:103415.

[21] Piepenburg O, Williams CH, Stemple DL, et al. DNA detection using recombination proteins. PLoS Biol. 2006;4(7):e204.

[22] Lobato IM, O'Sullivan CK. Recombinase polymerase amplification: basics, applications and recent advances. Trends Anal Chem. 2018;98:19–35.

[23] Ghouneimy A, Mahas A, Marsic T, et al. CRISPR-Based diagnostics: challenges and potential solutions toward point-of-care applications. ACS Synth Biol. 2023;12:1–16.

[24] Blanco-Miguez A, Beghini F, Cumbo F, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species with MetaPhlAn 4. 2022;2022.08.22.504593 Preprint at. DOI:10.1101/2022.08.22.504593

[25] Untergasser A, Cutcutache I, Koressaar T, et al. Primer3—new capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115. DOI:10.1093/nar/gks596

[26] Lorenz R, Bernhart SH, Höner Zu Siederdissen C, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011;6(1):26. DOI:10.1186/1748-7188-6-26

[27] Lemos J, Palmer SR, Zeng L, et al. The biology of Streptococcus mutans. Microbiol Spectr. 2019;7(1):1–7. DOI:10.1128/microbiolspec.GPP3-0051-2018

[28] Ajdić D, McShan WM, McLaughlin RE, et al. Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen. Proc Natl Acad Sci. 2002;99 (22):14434–14439. DOI:10.1073/pnas.172501299

[29] Downes J, Mantzourani M, Beighton D, et al. Scardovia wiggsiae sp. Nov., isolated from the human oral cavity and clinical material, and emended descriptions of the genus Scardovia and Scardovia inopinata. Int J Syst Evol Microbiol. 2011;61(1):25–29. DOI:10.1099/ijs.0.019752-0

[30] Oscarsson J, Claesson R, Lindholm M, et al. Tools of Aggregatibacter actinomycetemcomitans to evade the host response. J Clin Med. 2019;8:1079.

[31] Nedergaard S, Kobel CM, Nielsen MB, et al. Whole genome sequencing of aggregatibacter actinomyce-temcomitans cultured from blood stream infections reveals three major phylogenetic groups including a novel lineage expressing serotype a membrane O Polysaccharide. Pathogens. 2019;8(4):256. DOI:10.3390/pathogens8040256

[32] How KY, Song KP, Chan KG. Porphyromonas gingivalis: an overview of periodontopathic pathogen below the gum line. Front Microbiol. 2016;7:53.

[33] Chastain-Gross RP, Xie G, Bélanger M, et al. Genome sequence of porphyromonas gingivalis strain 381. Genome Announc. 2017;5(2):e01467-16. doi:10.1128/genomeA.01467-16.

[34] Zawadzki PJ, Perkowski K, Padzik M, et al. Examination of oral microbiota diversity in adults and older adults as an approach to prevent spread of risk factors for human infections. BioMed Res Int. 2017;2017:1–7.

[35] Traglia GM, Chua K, Centrón D, et al. Whole-genome sequence analysis of the naturally competent Acinetobacter baumannii clinical isolate A118. Genome Biol Evol. 2014;6:2235–2239.

[36] Botelho-Nevers E, Gouriet F, Lepidi H, et al. Chronic nasal infection caused by Klebsiella rhinoscleromatis or Klebsiella ozaenae: two forgotten infectious diseases. Int J Infect Dis. 2007;11(5):423–429. DOI:10.1016/j.ijid.2006.10.005

[37] Conlan S, Kong HH, Segre JA. Species-level analysis of DNA sequence data from the NIH human microbiome project. PLoS ONE. 2012;7:e47075.

[38] Rafiq Z, Sam N, Vaidyanathan R. Whole genome sequence of Klebsiella pneumoniae U25, a hypermu-coviscous, multidrug resistant, biofilm producing isolate from India. Mem Inst Oswaldo Cruz. 2016;111:144–146.

[39] McCormack M, Smith AJ, Akram AN, et al. Staphylococcus aureus and the oral cavity: an overlooked source of carriage and infection? Am J Infect Control. 2015;43(1):35–37. DOI:10.1016/j.ajic.2014.09.015

[40] Zubair S, Fischer A, Liljander A, et al. Complete genome sequence of Staphylococcus aureus, strain ILRI_Eymole1/1, isolated from a Kenyan dromedary camel. Stand Genomic Sci. 2015;10(1):109. DOI:10.1186/s40793-015-0098-6

[41] Chipashvili O, Utter DR, Bedree JK, et al. Episymbiotic Saccharibacteria suppresses gingival inflammation and bone loss in mice through host bacterial modulation. Cell Host Microbe. 2021;29:1649–1662.e7.

[42] Tavares LJ, Klein MI, Panariello BHD, et al. An in vitro model of Fusobacterium nucleatum and Porphyromonas gingivalis in single- and dual-species biofilms. J Periodontal Implant Sci. 2018;48:12.

[43] Kim S, Song M, Roh B-D, et al. Inhibition of Streptococcus mutans biofilm formation on composite resins containing ursolic acid. Restor Dent Endod. 2013;38:65–72.

[44] Arvikar SL, Hasturk H, Strle K, et al. Periodontal inflammation and distinct inflammatory profiles in saliva and gingival crevicular fluid compared with serum and joints in rheumatoid arthritis patients. J Periodontol. 2021;92:1379–1391.

[45] Escapa IF, Chen T, Huang Y, et al. New insights into human nostril microbiome from the Expanded Human Oral Microbiome Database (eHOMD): a resource for the microbiome of the human aerodigestive tract. mSystems. 2018;3(6):e00187-18. DOI:10.1128/mSystems.00187-18.

[46] Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581–583. DOI:10.1038/nmeth.3869

[47] Katoh K, Misawa K, Kuma K, et al. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30:3059–3066.

[48] Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–1973.

[49] Nguyen L-T, Schmidt HA, von Haeseler A, et al. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–274.

[50] Gut Microbiome Testing. Health supplements & probiotics. Viome.com. Available from: https://www.viome.com.

[51] Aas JA, Paster BJ, Stokes LN, et al. Defining the normal bacterial flora of the oral cavity. J Clin Microbiol. 2005;43:5721–5732.

[52] Eren AM, Borisy GG, Huse SM, et al. Oligotyping analysis of the human oral microbiome. Proc Natl Acad Sci. 2014;111:E2875–2884.

[53] Baumgartner D, Johannsen B, Specht M, et al. OralDisk: a chair-side compatible molecular platform using whole saliva for monitoring oral health at the dental practice. Biosensors (Basel). 2021;11(11):423. DOI:10.3390/bios11110423

[54] OralDNA Labs. Available from: https://www.oraldna.com/.

[55] Chandrasekaran SS, Agrawal S, Fanton A, et al. Rapid detection of SARS-CoV-2 RNA in saliva via Cas13. Nat Biomed Eng. 2022;6:944–956.

[56] Lee RA, Puig HD, Nguyen PQ, et al. Ultrasensitive CRISPR-based diagnostic for field-applicable detection of Plasmodium species in symptomatic and asymptomatic malaria. Proc Natl Acad Sci. 2020;117(41):25722–25731. DOI:10.1073/pnas.2010196117

[57] Notomi T, Okayama H, Masubuchi H, et al. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. 2000;28:e63.

[58] Maekawa T, Krauss J, Abe T, et al. Porphyromonas gingivalis manipulates complement and TLR signaling to uncouple bacterial clearance from inflammation and promote dysbiosis. Cell Host Microbe. 2014;15(6):768–778. DOI:10.1016/j.chom.2014.05.012

[59] Ackerman CM, Myhrvold C, Thakku SG, et al. Massively multiplexed nucleic acid detection with Cas13. Nature. 2020;582(7811):277–282. DOI:10.1038/s41586-020-2279-8