



Article

# How to Process Sputum Samples and Extract Bacterial DNA for Microbiota Analysis

Leonardo Terranova <sup>1,2,†</sup>, Martina Oriano <sup>3,4,†</sup>, Antonio Teri <sup>5</sup>, Luca Ruggiero <sup>1</sup>, Camilla Tafuro <sup>1</sup>, Paola Marchisio <sup>1</sup>, Andrea Gramegna <sup>3</sup>, Martina Contarini <sup>3</sup> , Elisa Franceschi <sup>3</sup> , Samantha Sottotetti <sup>5</sup>, Lisa Cariani <sup>5</sup>, Annamaria Bevivino <sup>6</sup> , James D. Chalmers <sup>7</sup>, Stefano Aliberti <sup>3,\*</sup> and Francesco Blasi <sup>3</sup>

<sup>1</sup> Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Pediatric Highly Intensive Care Unit, 20100 Milan, Italy; leonardo.terranova@policlinico.mi.it (L.T.); luca.ruggiero@policlinico.mi.it (L.R.); camilla.tafuro@policlinico.mi.it (C.T.); paola.marchisio@unimi.it (P.M.)

<sup>2</sup> Department of Clinical Sciences and Community Health, University of Milan, 20100 Milan, Italy

<sup>3</sup> Department of Pathophysiology and Transplantation, University of Milan, Internal Medicine Department, Respiratory unit and Adult Cystic Fibrosis Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, 20100 Milan, Italy; martina.orianio@unimi.it (M.O.); gramegna.med@gmail.com (A.G.); contarini.martina@gmail.com (M.C.); elisa.franceschi.93@gmail.com (E.F.); francesco.blasi@unimi.it (F.B.)

<sup>4</sup> Department of Molecular Medicine, University of Pavia, 27100 Pavia, Italy

<sup>5</sup> Cystic Fibrosis Microbiology Laboratory, Fondazione IRCCS Ca' Granda, 20100 Milan, Italy; antonioteri@hotmail.it (A.T.); samantha.sottotetti@hotmail.it (S.S.); lisa.cariani@gmail.com (L.C.)

<sup>6</sup> Department for Sustainability, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), 00123 Rome, Italy; annamaria.bevivino@gmail.com

<sup>7</sup> Scottish Centre for Respiratory Research, University of Dundee, Ninewells Hospital and Medical School, DD1 9SY Dundee, UK; j.chalmers@dundee.ac.uk

\* Correspondence: stefano.aliberti@unimi.it; Tel.: +39-0250320627 or +39-3394171538; Fax: +390250320625

† These two authors contributed equally and are considered first authors for this manuscript.

Received: 31 August 2018; Accepted: 13 October 2018; Published: 20 October 2018



**Abstract:** Different steps and conditions for DNA extraction for microbiota analysis in sputum have been reported in the literature. We aimed at testing both dithiothreitol (DTT) and enzymatic treatments of sputum samples and identifying the most suitable DNA extraction technique for the microbiota analysis of sputum. Sputum treatments with and without DTT were compared in terms of their median levels and the coefficient of variation between replicates of both DNA extraction yield and real-time PCR for the 16S rRNA gene. Treatments with and without lysozyme and lysostaphin were compared in terms of their median levels of real-time PCR for *S. aureus*. Two enzyme-based and three beads-based techniques for DNA extraction were compared in terms of their DNA extraction yield, real-time PCR for the 16S rRNA gene and microbiota analysis. DTT treatment decreased the coefficient of variation between replicates of both DNA extraction yield and real-time PCR. Lysostaphin (either 0.18 or 0.36 mg/mL) and lysozyme treatments increased *S. aureus* detection. One enzyme-based kit offered the highest DNA yield and 16S rRNA gene real-time PCR with no significant differences in terms of alpha-diversity indexes. A condition using both DTT and lysostaphin/lysozyme treatments along with an enzymatic kit seems to be preferred for the microbiota analysis of sputum samples.

**Keywords:** DNA extraction; sputum; microbiota; sequencing

## 1. Introduction

The study of the human microbiome has emerged as an important field in translational research over the past decade, and promising applications of this technique in clinical practice have been

suggested [1]. Most of the recent advances in microbial ecology have been made possible thanks to the development of next-generation sequencing technologies which allow the analysis of complex microbial populations in different body niches, including the respiratory tract. Multiple studies of the human-associated microbiota have suggested a fundamental role of microbes, and predominantly bacteria, in maintaining health and a role for changes in lung microbiota in contributing to diseases such as chronic obstructive pulmonary disease (COPD), bronchiectasis and cystic fibrosis (CF) [2–7].

Most of the published literature evaluating microbiota in chronic respiratory infections has used sputum as the preferred sample type. Although potentially less representative of the deep lung as compared with bronchoalveolar lavage (BAL), sputum could represent the preferred non-invasive diagnostic technique to be used in clinical practice for microbiota analysis. However, metagenomic sequencing from sputum is challenging because of the complex matrix of sputum, its heterogeneity, the activity of proteases and nucleases and the variability in sputum quality. Little has been published regarding the optimal methods for nucleic acid extraction from sputum for use in microbiome studies. Some bacteria require high levels of mechanical disruption or enzymatic disruption in order to achieve efficient lysis, and these include important pathogens in bronchiectasis and cystic fibrosis such as *Staphylococcus aureus*. [8,9].

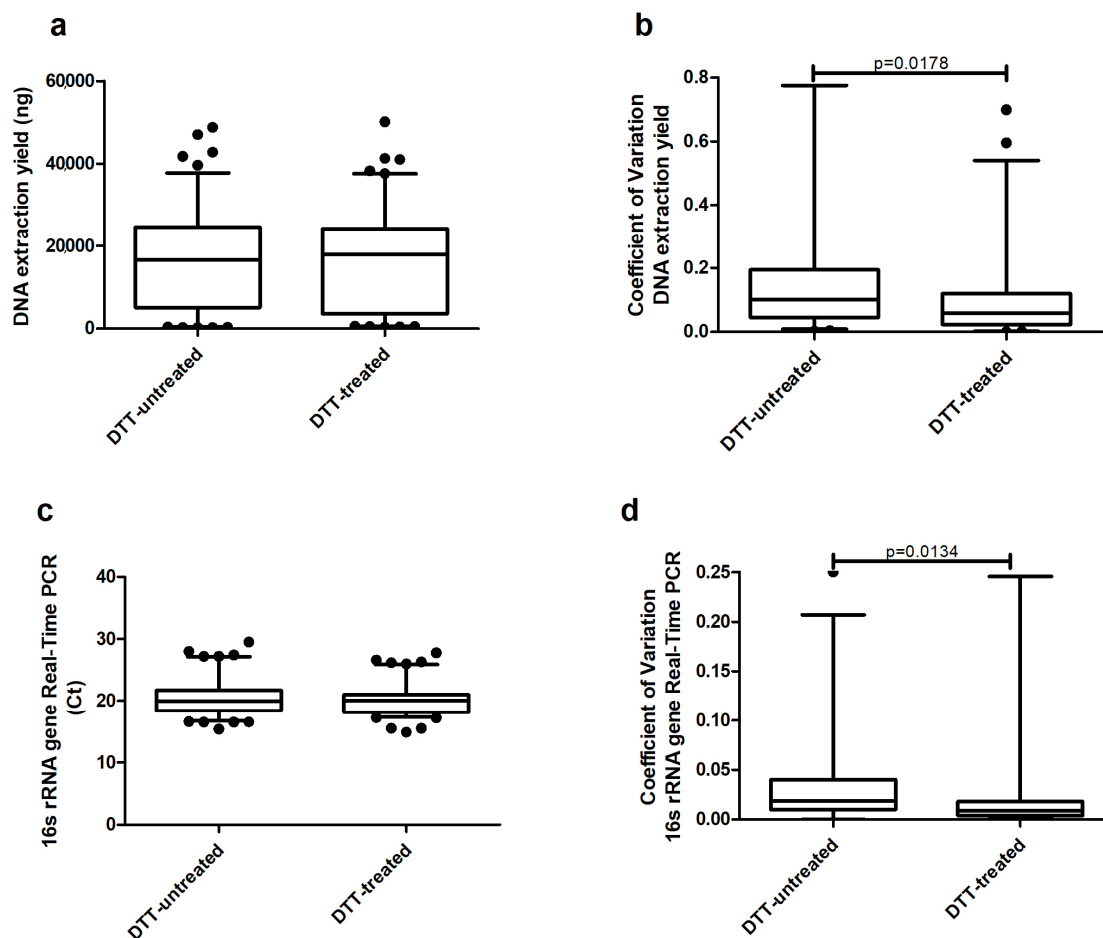
Different techniques have been used so far for microbiota analysis in respiratory samples, resulting in heterogeneity in terms of sample processing for DNA extraction, library preparation and bioinformatics analysis [4,10]. With a view to being able to compare the results of respiratory microbiota studies from different cohorts and different diseases, this heterogeneity needs to be approached, and optimal operating procedures to evaluate microbiota from sputum should be identified. Several steps should be standardized, including the use or not of fluidifying agents, such as dithiothreitol (DTT), or enzymes for specific bacterial lysis, such as lysozyme and lysostaphin for *S. aureus*, as well as different methods for DNA extraction and isolation. Although a comparison between different methods for DNA extraction in both BAL and gastrointestinal samples has already been carried out, no previous studies have specifically addressed this topic for sputum samples [11,12].

The present study had three objectives: (1) to evaluate the performance of DTT treatment to homogenize sputum samples; (2) to evaluate the effect of an enzymatic step for Gram-positive lysis; and (3) to investigate the performance of lytic chemical/enzyme-based and bead-based extraction kits.

## 2. Results

**Study patients.** A total of 32 adult patients with CF and bronchiectasis were enrolled in this study, and 77 sputum samples were collected for the three experiments.

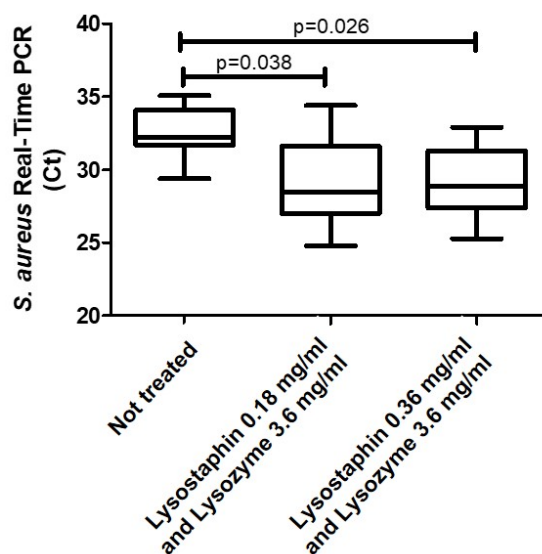
**Experiment 1—DTT treatment evaluation.** The median values of extracted DNA were similar between the DTT-untreated and DTT-treated group (16.630 vs. 17.880  $\mu\text{g}$ ,  $p = 0.7211$ ) (Figure 1a). The coefficient of variation between replicates was statistically lower in sputa treated with DTT in comparison with not-treated aliquots (DTT-untreated 0.1014 vs. DTT-treated: 0.0592,  $p = 0.018$ ), (Figure 1b). A similar situation has been identified when the analysis of the 16s real-time PCR was performed (Figure 1c,d). Although no differences were detected between groups in terms of their median Ct (DTT-untreated: 19.93 vs. DTT-treated: 20.09,  $p = 0.519$ ), there was a statistically significant decrease of the coefficient of variation in the DTT group compared to the not-treated one (DTT-untreated: 0.01859 vs. DTT-treated: 0.008848,  $p = 0.013$ ).



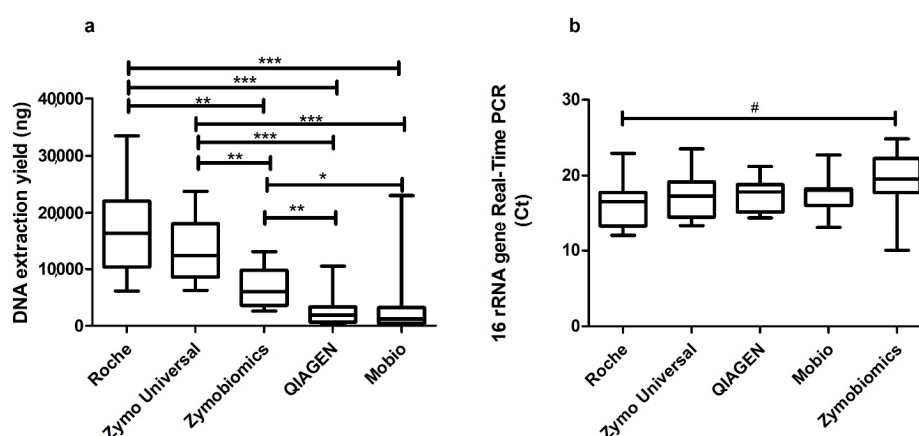
**Figure 1.** Comparison of median levels of (a) DNA extraction yield, (b) coefficient of variation between replicates, (c) 16s rRNA gene real-time PCR results, and (d) coefficient of variation between replicates across dithiothreitol (DTT)-untreated sputum vs. DTT-treated (Experiment 1).

**Experiment 2—Lytic enzymes evaluation.** *S. aureus* was detected from all the evaluated samples. Significantly high bacterial loads of *S. aureus*, represented by significantly lower median Ct levels by real-time polymerase chain reaction (PCR), were identified in samples treated with lytic enzymes compared with the non-treated ones ( $p = 0.038$  for lysostaphin 0.18 mg/mL and lysozyme 3.6 mg/mL;  $p = 0.027$  for lysostaphin 0.36 mg/mL and lysozyme 3.6 mg/mL). No significant difference was identified between the median levels of Ct among samples treated with the two dosages of lysostaphin and lysozyme ( $p = 1$ ) (Figure 2).

**Experiment 3—Evaluation of different kits for bacterial DNA extraction.** DNA extraction yield reached the highest values using the Roche kit in comparison to all the other kits. The median (interquartile range, IQR) yields for DNA extracted with the five kits were as follows: Roche 16,250 ng (6100–33,500 ng); Zymo Universal 12,350 ng (6225–23,750 ng); Zymobiomics 5975 ng (2625–13,000 ng); QIAGEN 1913 ng (204.5–10,475 ng); and Mobio 1205 ng (4.1–23,000 ng). Significant differences were found across the five kits ( $p < 0.0001$ ) and between two or more kits (Figure 3). No statistical differences were found between Roche and Zymo Universal in terms of DNA recovery ( $p = 0.3506$ ).



**Figure 2.** Comparison of median levels of real-time PCR for *S. aureus* across samples treated with lysostaphin at either 0.18 or 0.36 mg/mL and lysozyme (3.6 mg/mL) or without lytic enzymes treatment. Results expressed in cycle threshold (Ct). Kruskal–Wallis test:  $p = 0.041$  (Experiment 2).



**Figure 3.** Comparison of median levels of (a) DNA extraction yield and (b) bacterial extraction efficiency across the five analyzed commercial kits. Kruskal–Wallis test: DNA extraction  $p < 0.0001$ ; 16s rRNA gene real-time PCR  $p = 0.169$ . Mann–Whitney test: \*:  $p \leq 0.001$ ; \*\*:  $p \leq 0.0009$ , \*\*\*  $p \leq 0.0001$ ; #:  $p = 0.027$  (Experiment 3).

Results from the 16s rRNA gene Real-Time PCR showed a similar median Ct across the five extraction methods ( $p = 0.169$ ). The median (IQR) Ct values were as follows: Roche 16.57 (12.01–22.94); Zymo Universal 17.29 (13.42–23.55); QIAGEN 17.86 (14.45–21.22); Mobio 18.87 (13.16–22.72); and Zymobiomics 19.57 (10.04–24.85). The only significant difference was found between Ct values obtained using the Roche and Zymobiomics kits ( $p = 0.027$ ).

A low number of joined reads (min 47–max 253) was produced for each negative control and no bacteria were identified. None of the operational taxonomic units (OTUs) assigned to these samples passed the 0.5% prevalence threshold that was set to filter out low-abundance OTUs from the OTU table considered to be due to cross-talk or experimental error.

The sequencing depth was nearly the same for all samples which were extracted with the different kits (typical number of reads comprised between 0 and 100,000), except for one. One sample extracted with Zymobiomics was sequenced ~7 times more deeply than all the others (Figure 4).

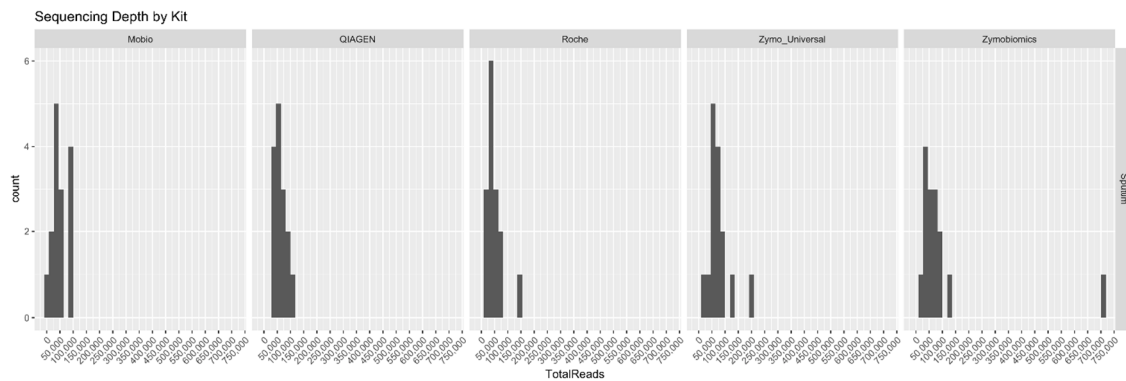


Figure 4. Sequencing depth of samples by kit.

The only sample we were not able to sequence belonged to the Mobio group.

None of the tested techniques showed a statistical difference compared to the others in any of the considered alpha indices from genera (Figure 5).

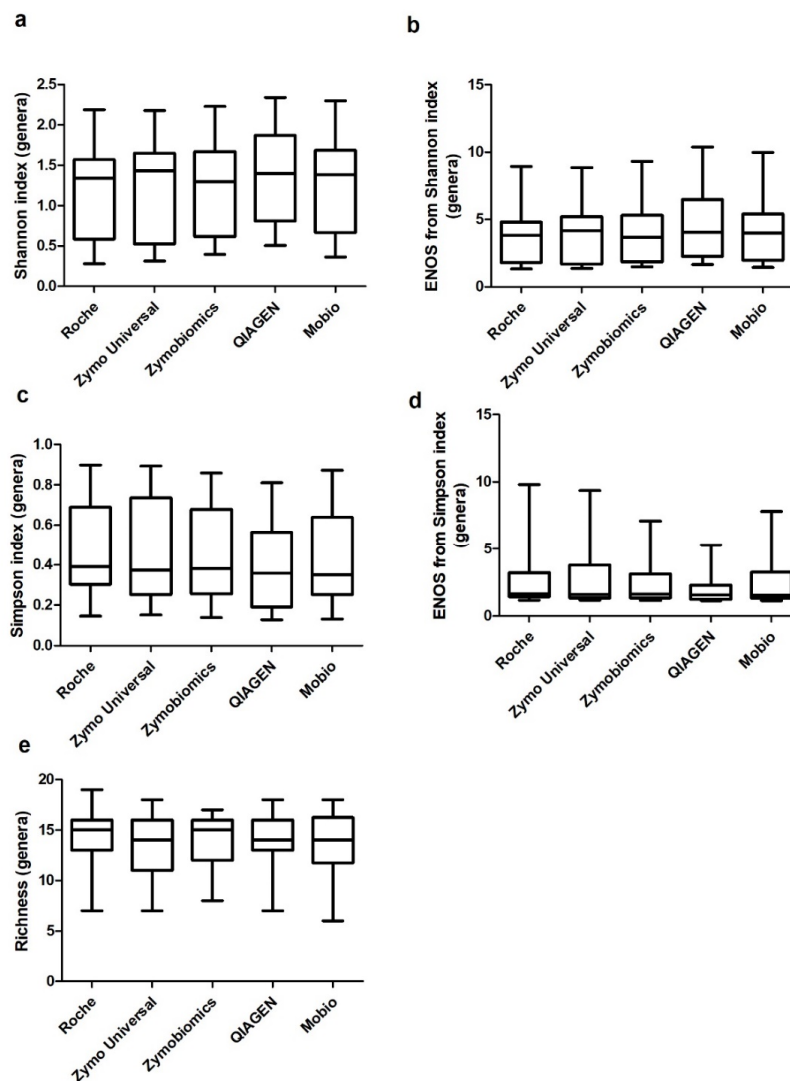


Figure 5. Comparison of median levels of (a) the Shannon index, (b) effective number of species (ENOS) from Shannon, (c) the Simpson index, (d) ENOS from the Simpson index, and (e) richness across the five analyzed commercial kits (Experiment 3).



The relative abundances of the 15 samples were measured using the five evaluated kits and are presented in Figure 6 along with the results of cultures from standard microbiology. The comparison between genera retrieved through 16S rRNA gene sequencing and culture results showed a complete concordance of the two techniques at the genus level, even if, as expected, 16S rRNA sequencing allowed us to identify a broader range of bacteria. Each technique seems to be able to extract bacterial genera in a similar way.



Figure 6. Genera-relative abundances of 15 patients for each of the considered methods (Experiment 3).

### 3. Discussion

Our study indicates that the High Pure PCR Template Preparation (Roche) kit preceded by both DTT and enzymatic digestion steps offers a higher DNA yield and more 16S rRNA gene real-time PCR extracted from sputum samples in comparison to the other evaluated conditions (Figure 3). Furthermore, no significant differences in terms of alpha-diversity (i.e., Shannon, Simpson and richness indices) have been detected across the different methods evaluated (Figure 5).

A previous pilot study showed that no negative effect derived from the use of enzymes, DTT and different kits was identified. Variables were evaluated simultaneously, but a single variable approach is needed in order to have a deep understanding of single variables' contributions [13]. For this reason, a consecutive approach has been adopted for the present study in order to understand the unique contribution of three different interventions: the use of DTT, the introduction of an enzymatic step and the choice of the best DNA extraction kit.

The reactive sulphhydryl groups of DTT have a mucolytic function and are used to reduce mucoprotein disulfide bonds. DTT is commonly used in standard microbiology to homogenize sputum samples, enhancing the isolation of microorganisms, and has been also adopted in different studies for microbiota analysis [14,15]. Sputum heterogeneity is one of the most challenging aspects of bacterial DNA recovery, especially in patients with CF and bronchiectasis. We evaluated the use of DTT through 16s rRNA real-time PCR in terms of (1) median levels, in order to understand if the presence of DTT was increasing bacterial DNA recovery; and (2) the coefficient of variation between replicates, in order to understand if the presence of DTT was able to better homogenize sputum and increase data reproducibility (Figure 1). The use of a reducing agent to treat this matrix allows the release of those bacteria entrapped in sputum plugs. We demonstrated that the addition of DTT increases the reproducibility of the data and facilitates sputum handling without affecting DNA extraction.

Lysostaphin and lysozyme have been extensively used in the literature to increase DNA recovery from Gram-positive bacteria [16–18]. These enzymes specifically lyse cell walls and ensure DNA recovery. Previously published experiences demonstrated clear advantages in adding lytic enzymes, including lysozyme and lysostaphin, to sputum samples. Cell wall destruction determined by the addition of these products improves real-time PCR detection and the sequencing of some Gram-positive bacteria, such as *S. aureus*, and this is evident already at low enzyme concentrations (lysostaphin at 0.18 mg/mL) [5]. Williamson and colleagues demonstrated the need to introduce these agents for microbiome analysis, even if there is no agreement in the literature on the concentration of lysostaphin needed to lyse bacteria in sputum samples. We demonstrated that sputum processing using these enzymes increases *S. aureus* DNA recovery, while the increase of lysostaphin concentrations seems not to affect *S. aureus* detection (Figure 2). Even if the presence of both pre-treatments will increase the time and costs of bacterial DNA extraction from sputum, the addition of this step seems to be needed in order to have more reliable and reproducible data.

We decided to use the Zymo Universal kit according to a previously performed pilot study [13] which compared different commercial kits across all conditions (with and without DTT and enzymes) in order to understand if combining/interfering effects were present among different variables. In this instance, we found no combined effect derived from the use of enzymes and DTT in terms of DNA yield, 16s rRNA gene and the *S. aureus* real-time PCR cycle threshold and microbiota analysis.

Studies evaluating respiratory microbiota in COPD and bronchiectasis patients used either BAL, spontaneous sputum or induced sputum [19–22]. DNA extraction methods across these experiences have been very heterogeneous, with some authors using chemical methods [14], others bead beating methods [23] and others a combination of physical lysis and methods such as phenol-chloroform [22]. The use of a combination of chemical and enzyme-based lysis increases DNA recovery from sputum samples. This method reduces the possibility of losing material after lysis because the whole of the lysate from samples is transferred into the DNA purification column. When bead-beating techniques are used, a fraction of lysate is retained in the beads phase after centrifugation. Even if Mobio, Zymobiomics and QIAGEN are based on mechanical disruption through bead beating, differences in methods could explain differences in extracted DNA yield. As reported in Table 1, Zymobiomics and QIAGEN also expose samples to enzymatic digestion in order to increase DNA extraction. Moreover, beads sizes differ between kits: QIAGEN and Mobio present glass beads of 0.1 mm of size, while Zymobiomics has beads of both 0.1 and 0.5 mm of size. Finally, column filters and reagents are specific for each kit. In terms of microbiota analysis, we were not able to identify statistically significant

differences across the used kits, and we can therefore speculate that all the evaluated kits perform equally in representing analyzed samples (Figures 5 and 6).

No similar studies have been previously designed and published to evaluate different DNA extraction methods from sputum to microbiota evaluations. Because of this, we were not able to define the correct sample size of our study a priori which should be interpreted as a pilot in its nature. Moreover, samples were tested in duplicate in real-time PCR (not triplicate) due to the low DNA yield of some samples. In the present study, we were not able to evaluate other methods for DNA extraction, including boiling, phenol-chloroform or automatic extraction, which have been previously reported in the literature, although the ones we have considered are those commercially available [14,22].

This is the first study able to evaluate different steps for DNA extraction from sputum for microbiota analysis in a sequential process which permits us to better control the impact of different interventions. The final evaluation consisted of the microbiota analysis of all the possible final conditions in order to have a general overview of the method. The use of blanks from each step and a positive control to increase the reliability of the sequencing techniques allowed us to exclude kits and environmental bacterial contamination.

#### 4. Materials and Methods

Three sequential experiments were performed to specifically address the study objectives. Sputum samples were collected from 32 adults with either CF or bronchiectasis referring to the Respiratory Department of Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy, between February and March 2018. Subjects signed an informed consent form and gave their approval to use their samples for the purpose of this study. Aliquots of sputum were collected and stored at  $-80^{\circ}\text{C}$  for all subsequent analyses, after which extraction samples were thawed and 0.1 g of sputum plug was selected for extractions.

Standard cultures for bacteria, fungi and non-tuberculous mycobacteria were performed by our local Cystic Fibrosis laboratory in accordance with the CF Foundation guidelines [24]. 100 microliters of the samples were plated on complete and selective media to isolate all potential pathogenic microorganisms. In the case of growth of potential pathogens, the colony morphotypes observed on the selective and non-selective media were identified by colonial morphology, pigment production or  $\beta$ -haemolysis, and then by biochemical assays and/or proteomic profiling by matrix assisted laser desorption-time of flight mass spectrometry (MALDI-TOF MS). After that, antibiotic susceptibility tests were performed by the Microscan WalkAway plus System (Beckman Coulter, Brea, CA, USA) [15].

**Experiment 1—DTT Treatment Evaluation.** In order to evaluate the need for sample homogenization, a total of 10 patients were enrolled in this experiment and 55 sputum samples were collected. Each sample was divided into two aliquots: the first aliquot was treated 1:1 with DTT (Sputafluid, Biolife, Italy) following the manufacturer's instructions (DTT solution 10%, vortexed and left at room temperature for 15 min), while the second one received no treatment. After that, DNA was extracted in duplicate using Zymo Quick-DNA Universal Kit (Zymo, Irvine, CA, USA) according to the manufacturer's instructions and eluted in 50  $\mu\text{L}$  elution buffer. DNA extraction yield was measured through quantification by the Quant-iT dsDNA Assay Kit High Sensitivity and Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). Subsequently, samples were diluted at 5 ng/ $\mu\text{L}$  to avoid the inhibition of PCR due to the high concentration of human DNA and tested through real-time PCR for SYBR Green for 16S rRNA gene amplification [25]. Each sample was tested in duplicate and the cycle threshold (Ct) means between replicates, standard deviation and coefficient of variation were considered in order to identify potential differences between replicates. Endpoints for this experiment were real-time PCR for the 16S rRNA gene and coefficient of variation, as well as DNA yield and coefficient of variation. The DNA yield and 16S rRNA gene real-time PCR of the two groups identified as DTT-untreated and DTT are presented as a median with interquartile range (IQR). The coefficient of variation as the ratio of the standard deviation to the mean was calculated. The Mann-Whitney test was applied between groups.



**Experiment 2—Lytic enzymes evaluation.** In order to understand if lysis using lysostaphin and lysozyme would increase sensitivity for specific genera, sputum samples from seven patients in which *S. aureus* had been isolated according to standard microbiology were selected and treated with DTT according to the results of Experiment 1. The enzymatic digestion was inserted in the procedure immediately after DTT treatment, carried out with lysozyme at 3.6 mg/mL and lysostaphin in two different conditions: 0.18 and 0.36 mg/mL (Sigma-Aldrich, Saint Louis, MO, USA). Samples were incubated at 37 °C for 30 min prior to DNA extraction. DNA extraction and DTT treatment were carried out as reported above. Real-time PCR for *S. aureus* was conducted on DNA extracts as already described [26]. Data were divided into three groups: (i) no lysostaphin; (ii) lysostaphin 0.18 mg/mL and (iii) lysostaphin 0.36 mg/mL, which are presented as medians (IQR). The endpoint included real-time PCR for *S. aureus*. Mann–Whitney and Kruskal–Wallis tests were applied to the three groups.

**Experiment 3—Evaluation of different kits for bacterial DNA extraction.** Fifteen samples were analyzed to understand which is the most suitable technique to extract DNA from sputum. Samples were pretreated according to Experiment 1 and 2 (DTT treatment and lysostaphin 0.18 mg/mL and lysozyme 3.6 mg/mL) and DNA was extracted using five different commercial kits: (1) Roche High Pure PCR Template Preparation Kit (Hoffmann, La Roche, Basel, Switzerland); (2) Zymo Quick-DNA Universal Kit (Zymo, Irvine, CA, USA); (3) MoBio PowerLyzer PowerSoil DNA isolation kit (Mobio, Loker Ave West, Carlsbad, CA, USA), actually sold by QIAGEN as the Qiagen DNeasy PowerSoil kit; (4) QIAGEN QIAmp Cador Pathogen Mini kit (Qiagen, Hilden, Germany); (5) ZymoBIOMICS DNA Miniprep Kit (Zymo, Irvine, CA, USA) (Table 1). The first two methods use the combination of both chemical and enzyme-based lysis, while the other three use a mechanical destruction through bead-beating. Commercial kits were used according to the manufacturer's instructions and DNA was eluted in 50 µL elution buffer. DNA extraction was carried out as previously described, as well as real-time PCR targeting the 16S rRNA gene. The microbiota evaluation procedure is reported below. Negative controls from each DNA extraction kit and from PCR were also sequenced in order to also evaluate environmental contamination in kit reagents.

**Microbiota evaluation.** The V3-V4 variable regions of the 16S rRNA gene were amplified from DNA extracts using the 16S metagenomic sequencing library preparation protocol (Illumina, San Diego, CA, USA). PCR products, approximately sized 630 base pairs, were visualized using microfluidics-based gel electrophoresis on Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA) and then were cleaned using AMPure XP magnetic bead-based purification (Beckman Coulter, Brea, CA, USA). Sample libraries were quantified using the Qubit as reported above and then pooled in an equimolar mode. Finally, the pool was sequenced on the MiSeq (Illumina, San Diego, CA, USA) sequencing platform, using a 2 × 300 cycle V3 kit and following standard Illumina sequencing protocols.

**Bioinformatic and statistical analyses.** Demultiplexed paired-end reads in the FASTQ format were received from the Illumina MiSeq instrument. Sequencing data were processed following the UPARSE pipeline by Edgar [27], using USEARCH v10.0.240 (Tiburon, CA, USA) [28] and VSEARCH v2.3.4 (Oslo, Norway) [29]. Overall run quality was checked using FastQC v0.11.2 (Cambridge, UK) [30] and reports were summarized using MultiQC v1.4 (Stockholm, Sweden) [31]. Quality scores dropped towards the end of the reverse reads, so they were globally trimmed at position 275 before merging with the corresponding forward reads. Parameters for paired-end reads merging were set as follows: a minimum overlapping length of 95 base pairs, a minimum 90% identity of alignment, and the merged sequence lengths were restricted to 432–482 bases. Consensus sequences from all samples were pooled together and primers were stripped from both ends. This “raw” set of merged sequences was then quality-filtered and de-replicated to obtain a subset of high-quality unique sequences to be clustered into operational taxonomic units (OTUs). Sequences with more than one expected number of errors (EE) were discarded and singletons removed during de-replication. OTUs were clustered at a 97% identity threshold. The taxonomy prediction at the genus level for OTU sequences was performed via the SINTAX algorithm [32], using the RDP training set v16 as the reference database and 0.8 as confidence threshold. An OTU table was constructed by mapping the whole set of “raw” merged

paired-end reads to the representative set of OTUs, using a 97% identity threshold. It was then filtered for low-abundance OTUs (<0.5 overall frequency), which were discarded, and normalized to the same number of reads per sample. This OTU table was used for all downstream analyses. Alpha diversity was measured for each sample using different metrics (Shannon entropy, Simpson estimators, and richness). These indices were then converted to the effective number of species (ENOS) [33] to be easily compared to each other. Results at the genus level were considered. Data were divided in five groups depending on the DNA extraction technique and presented as median (IQR). Sequences have been uploaded to the Sequence Read Archive (SRA) and are available under the project ID PRJNA488913. Endpoints included the DNA yield, real-time PCR for the 16S rRNA gene and alpha indices (Shannon and Simpson expressed as indices and as ENOS, and richness). Furthermore, a relative abundances comparison with standard microbiology was conducted. Mann–Whitney and Kruskal–Wallis tests were applied. A significance level of 0.05 was used for all the three experiments using GraphPad Prism 5 (La Jolla, CA, USA).

**Table 1.** Comparison of different methods for DNA extraction (Information provided by manufacturer’s instructions and protocols).

Method	Basis of Function	Cost per Sample	Time per Sample (min)	Advantages	Disadvantages	Reference
Roche High Pure PCR Template Preparation Kit	Lysis Buffer and proteinase K digestion	2.62 €	16	Easy to use and quick protocol.	Milder extraction method than mechanical	Feigelman et al. 2017 [34]
Zymo Quick-DNA Universal Kit	Lysis Buffer and proteinase K digestion	2.49 €	18	Easy to use and quick protocol	Milder extraction method than mechanical	Blow et al. 2017 [35]
MoBio PowerLyzer PowerSoil DNA isolation kit	Bead beating	6.16 €	41	Possibility to lyse hard-to-lyse bacteria	Time consuming, possibility of DNA loss because of high number of steps	Willner et al. 2012 [36]
QIAGEN QIAmp Cador Pathogen Mini kit	Bead beating and proteinase K digestion	4.6 €	20	Possibility to lyse hard-to-lyse bacteria. Possibility to combine different methods.	Possibility of DNA loss	Hart et al. 2015 [37]
ZymoBIOMICS DNA Miniprep Kit	Bead beating and proteinase K digestion	5.12 €	63	Combination of enzymatic and mechanical disruption. Option to use different protocols	Time consuming, possibility of DNA loss because of high number of steps	Sohrabi et al. 2016 [38]

## 5. Conclusions

In conclusion, the use of the Roche kit with the addition of a DTT treatment and enzymatic digestion with lysostaphin (0.18 mg/mL) and lysozyme offers better performance in terms of DNA yield and 16S rRNA gene real-time PCR extracted from sputum samples in comparison to other conditions.

**Author Contributions:** Conceptualization, S.A., L.T. and M.O.; methodology, L.T., M.O. and L.R.; validation, L.T., M.O. and L.R.; formal analysis, C.T. and M.O.; investigation, S.A., L.T. and M.O.; resources, F.B., P.M., S.A., L.C., A.T., A.G., M.C. and E.F.; data curation, C.T.; writing—original draft preparation, S.A., L.T. and M.O.; writing—review & editing, L.R., C.T., P.M., A.G., M.C., E.F., A.T., S.S., L.C., A.B., J.D.C. and F.B.; supervision, F.B., S.A. and P.M.; project administration, S.A.; funding acquisition, F.B., S.A. and P.M.

**Acknowledgments:** The authors acknowledge the support of all the healthcare professionals involved in the Bronchiectasis and PCD Program and the Adult Cystic Fibrosis Program of the Policlinico Hospital in Milan

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Rogers, G.B.; Shaw, D.; Marsh, R.L.; Carroll, M.P.; Serisier, D.J.; Bruce, K.D. Respiratory microbiota: Addressing clinical questions, informing clinical practice. *Thorax* **2015**, *70*, 74–81. [[CrossRef](#)] [[PubMed](#)]
2. Muhlebach, M.S.; Zorn, B.T.; Esther, C.R.; Hatch, J.E.; Murray, C.P.; Turkovic, L.; Ranganathan, S.C.; Boucher, R.C.; Stick, S.M.; Wolfgang, M.C. Initial acquisition and succession of the cystic fibrosis lung microbiome is associated with disease progression in infants and preschool children. *PLoS Pathog.* **2018**, *14*, e1006798. [[CrossRef](#)] [[PubMed](#)]

3. Ubags, N.D.J.; Marsland, B.J. Mechanistic insight into the function of the microbiome in lung diseases. *Eur. Respir. J.* **2017**, *50*, 1602467. [[CrossRef](#)] [[PubMed](#)]
4. Tunney, M.M.; Einarsson, G.G.; Wei, L.; Drain, M.; Klem, E.R.; Cardwell, C.; Ennis, M.; Boucher, R.C.; Wolfgang, M.C.; Elborn, S.E. Lung microbiota and bacterial abundance in patients with bronchiectasis when clinically stable and during exacerbation. *Am. J. Respir. Crit. Care Med.* **2013**, *187*, 1118–1126. [[CrossRef](#)] [[PubMed](#)]
5. Williamson, K.M.; Wagner, B.D.; Robertson, C.E.; Johnson, E.J.; Zemanick, E.T.; Kirk Harris, J. Impact of enzymatic digestion on bacterial community composition in CF airway samples. *PeerJ* **2017**, *5*, e3362. [[CrossRef](#)] [[PubMed](#)]
6. Pragman, A.A.; Lyu, T.; Baller, J.A.; Gould, T.J.; Kelly, R.F.; Reilly, C.S.; Isaacson, R.E.; Wendt, C.H. The lung tissue microbiota of mild and moderate chronic obstructive pulmonary disease. *Microbiome* **2018**, *6*, 7. [[CrossRef](#)] [[PubMed](#)]
7. Bacci, G.; Paganin, P.; Lopez, L.; Vanni, C.; Dalmastrì, C.; Daddiego, L.; Perrotta, G.; Dolce, D.; Morelli, P.; Tuccio, V.; et al. Pyrosequencing Unveils Cystic Fibrosis Lung Microbiome Differences Associated with a Severe Lung Function Decline. *PLoS ONE* **2016**, *11*, e0156807. [[CrossRef](#)]
8. Parkins, M.D.; Floto, R.A. Emerging bacterial pathogens and changing concepts of bacterial pathogenesis in cystic fibrosis. *J. Cyst. Fibros.* **2015**, *14*, 293–304. [[CrossRef](#)] [[PubMed](#)]
9. King, P.T.; Holdsworth, S.R.; Freezer, N.J.; Villanueva, E.; Holmes, P.W. Microbiologic follow-up study in adult bronchiectasis. *Respir. Med.* **2007**, *101*, 1633–1638. [[CrossRef](#)] [[PubMed](#)]
10. Hogan, D.A.; Willger, S.D.; Dolben, E.L.; Hampton, T.H.; Stanton, B.A.; Morrison, H.G.; Sogin, M.L.; Czum, J.; Ashare, A. Analysis of lung microbiota in bronchoalveolar lavage, protected brush and sputum samples from subjects with mild-to- moderate cystic fibrosis lung disease. *PLoS ONE* **2016**, *11*, e0149998. [[CrossRef](#)] [[PubMed](#)]
11. Lim, M.Y.; Song, E.J.; Kim, S.H.; Lee, J.; Nam, Y.D. Comparison of DNA extraction methods for human gut microbial community profiling. *Syst. Appl. Microbiol.* **2018**, *41*, 151–157. [[CrossRef](#)] [[PubMed](#)]
12. Wen, Y.; Xiao, F.; Wang, C.; Wang, Z. The impact of different methods of DNA extraction on microbial community measures of BALF samples based on metagenomic data. *Am. J. Transl. Res.* **2016**, *8*, 1412–1425. [[PubMed](#)]
13. Oriano, M.; Terranova, L.; Ruggiero, L.; Tafuro, C.; Teri, A.; Franceschi, E.; Amati, F.; Gramegna, A.; Contarini, M.; Cariani, L.; et al. What Is the Best Technique to Extract Bacterial DNA from Sputum? ERS International Congress: Paris, France, 2018.
14. Dicker, A.J.; Crichton, M.L.; Pumphrey, E.G.; Cassidy, A.J.; Suarez-Cuartin, G.; Sibila, O.; Furrie, E.; Fong, C.J.; Ibrahim, W.; Brady, G.; et al. Neutrophil extracellular traps are associated with disease severity and microbiota diversity in patients with chronic obstructive pulmonary disease. *J. Allergy. Clin. Immunol.* **2018**, *141*, 117–127. [[CrossRef](#)] [[PubMed](#)]
15. Denton, M.; Doherty, C.; Foweraker, J.; Govan, J.; Hall, M.; Isalska, B.; Jones, A.; Kenna, D.; Larsen, A.; Pitt, T.; et al. Laboratory standards for processing microbiological samples from people with cystic fibrosis. In *Cystic Fibrosis Trust; Report of the UK Cystic Fibrosis Trust Microbiology Laboratory Standards Working Group*: Kent, UK, 2010.
16. Zhao, J.; Carmody, L.A.; Kalikin, L.M.; Li, J.; Petrosino, J.F.; Schloss, P.D.; Young, V.B.; LiPuma, J.J. Impact of Enhanced Staphylococcus DNA extraction on microbial community measures in cystic fibrosis sputum. *PLoS ONE* **2012**, *7*, e33127. [[CrossRef](#)] [[PubMed](#)]
17. Zhou, Y.; Lin, P.; Li, Q.; Han, L.; Zheng, H.; Wei, Y.; Cui, Z.; Ni, Y.; Guo, X. Analysis of the microbiota of sputum samples from patients with lower respiratory tract infections. *Acta Biochim. Biophys. Sin.* **2010**, *42*, 754–761. [[CrossRef](#)] [[PubMed](#)]
18. Chen, C.; Shen, T.; Tian, F.; Lin, P.; Li, Q.; Cui, Z.; Zhang, Y.; Xue, M.; Ye, J.; Guo, X.; et al. New microbiota found in sputum from patients with community-acquired pneumonia. *Acta Biochim. Biophys. Sin.* **2013**, *45*, 1039–1048. [[CrossRef](#)] [[PubMed](#)]
19. Dickson, R.P.; Erb-Downward, J.R.; Prescott, H.C.; Martinez, F.J.; Curtis, J.L.; Lama, V.N.; Huffnagle, G.B. Cell-associated bacteria in the human lung microbiome. *Microbiome* **2014**, *2*, 28. [[CrossRef](#)] [[PubMed](#)]
20. Rogers, G.B.; van der Gast, C.J.; Cuthbertson, L.; Thomson, S.K.; Bruce, K.D.; Martin, M.L.; Serisier, D.J. Clinical measures of disease in adult non-CF bronchiectasis correlate with airway microbiota composition. *Thorax* **2013**, *68*, 731–737. [[CrossRef](#)] [[PubMed](#)]

21. Cox, M.J.; Turek, E.M.; Hennessy, C.; Mirza, G.; James, P.L.; Coleman, M.; Jones, A.; Wilson, R.; Bilton, D.; Cookson, W.O.C.; et al. Longitudinal assessment of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis bronchiectasis patients. *PLoS ONE* **2017**, *12*, e0170622. [CrossRef] [PubMed]
22. Rogers, G.B.; Zain, N.M.M.; Bruce, K.D.; Burr, L.D.; Chen, A.C.; Rivett, D.W.; McGuckin, M.A.; Serisier, D.J. A novel microbiota stratification system predicts future exacerbations in bronchiectasis. *Ann. Am. Thorac. Soc.* **2014**, *11*, 496–503. [CrossRef] [PubMed]
23. Purcell, P.; Jary, H.; Perry, A.; Perry, J.D.; Stewart, C.J.; Nelson, A.; Lanyon, C.; Smith, D.L.; Cumming, S.P.; De Soya, A. Polymicrobial airway bacterial communities in adult bronchiectasis patients. *BMC Microbiol.* **2014**, *14*, 130. [CrossRef] [PubMed]
24. Farrell, P.M.; White, T.B.; Ren, C.; Hempstead, S.E.; Accurso, F.; Derichs, N.; Howenstine, M.; McColley, S.A.; Rock, M.; Rosenfeld, M.; et al. Diagnosis of Cystic Fibrosis: Consensus Guidelines from the Cystic Fibrosis Foundation. *J. Pediatr.* **2018**, *181*, S4–S15.e1. [CrossRef] [PubMed]
25. Vega, M.F.; Dieguez, S.N.; Riccio, B.; Aranguren, S.; Giordano, A.; Denzoin, L.; Soraci, A.L.; Tapia, M.O.; Ross, R.; Apás, A.; et al. Zearalenone adsorption capacity of lactic acid bacteria isolated from pigs. *Braz. J. Microbiol.* **2017**, *48*, 715–723. [CrossRef] [PubMed]
26. Gadsby, N.J.; McHugh, M.P.; Russell, C.D.; Mark, H.; Morris, A.C.; Laurenson, I.F.; Hill, A.T.; Templeton, K.E. Development of two real-time multiplex PCR assays for the detection and quantification of eight key bacterial pathogens in lower respiratory tract infections. *Clin. Microbiol. Infect.* **2015**, *21*, 788.e1–788.e13. [CrossRef] [PubMed]
27. Edgar, R.C. UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **2013**, *10*, 996–998. [CrossRef] [PubMed]
28. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [CrossRef] [PubMed]
29. Rognes, T.; Flouri, T.; Nichols, B.; Quince, C.; Mahé, F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **2016**, *4*, e2584. [CrossRef] [PubMed]
30. Andrews, S. FastQC: A Quality Control Tool for High throughput Sequence Data. 2010. Available online: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (accessed on 5 February 2018).
31. Ewels, P.; Ankenbrand, M.J.; Hammarén, R.; Saveliev, V.; Booth, T.; Jones, A.; Neumann, T.; Magnusson, M.; Alexanderscholz; Andeer, R.; et al. MultiQC Version 1.4. 2016. Available online: <https://zenodo.org/record/1145834#.W2FsDi1aZBw> (accessed on 5 February 2018).
32. Edgar, R.C. SINTAX: A Simple Non-BAYESIAN Taxonomy Classifier for 16S and ITS Sequences. *bioRxiv* 2016. Available online: <https://www.biorxiv.org/content/early/2016/09/09/074161> (accessed on 4 July 2017).
33. Chao, A.; Chiu, C.-H.; Jost, L. Phylogenetic diversity measures based on Hill numbers. *Philos. Trans. R Soc. B Biol. Sci.* **2010**, *365*, 3599–3609. [CrossRef] [PubMed]
34. Feigelman, R.; Kahlert, C.R.; Baty, F.; Rassouli, F.; Kleiner, R.L.; Kohler, P.; Brutsche, M.H.; von Mering, C. Sputum DNA sequencing in cystic fibrosis: Non-invasive access to the lung microbiome and to pathogen details. *Microbiome* **2017**, *5*, 1–14. [CrossRef] [PubMed]
35. Blow, F.; Vontas, J.; Darby, A.C. Draft Genome Sequence of *Chryseobacterium* Strain CBo1 Isolated from *Bactrocera oleae*. *Genome Announc.* **2017**, *5*, e00177-17. [CrossRef] [PubMed]
36. Willner, D.; Daly, J.; Whiley, D.; Grimwood, K.; Wainwright, C.E.; Hugenholtz, P. Comparison of DNA extraction methods for microbial community profiling with an application to pediatric bronchoalveolar lavage samples. *PLoS ONE* **2012**, *7*, e34605. [CrossRef] [PubMed]
37. Hart, M.L.; Meyer, A.; Johnson, P.J.; Ericsson, A.C. Comparative Evaluation of DNA Extraction Methods from Feces of Multiple Host Species for Downstream Next-Generation Sequencing. *PLoS ONE* **2015**, *10*, e0143334. [CrossRef] [PubMed]
38. Sohrabi, M.; Nair, R.G.; Samaranayake, L.P.; Zhang, L.; Zulfiker, A.H.M.; Ahmetagic, A.; Good, D.; Wei, M.Q. The yield and quality of cellular and bacterial DNA extracts from human oral rinse samples are variably affected by the cell lysis methodology. *J. Microbiol. Methods* **2016**, *122*, 64–72. [CrossRef] [PubMed]

